GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 3, 2004, 12:11:47 ; Search time 11 Seconds
(without alignments)
23.668 Million cell updates/sec

Title: US-09-871-974-2
Perfect score: 29
Sequence: 1 TKPPR 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	99	1	D127_HUMAN	Q9h1m4 homo sapien
2	29	100.0	201	1	COAE_SYNY3	Q55515 synechocyst
3	29	100.0	227	1	FA3C_HUMAN	Q92520 homo sapien
4	29	100.0	227	1	FA3C_MOUSE	Q91vu0 mus musculu
5	29	100.0	262	1	RPC_BP163	P15238 bacteriopha
6	29	100.0	465	1	ENGA_STRCO	Q9eww8 streptomyce
7	29	100.0	491	1	ENGA_STRAW	Q828y7 streptomyce
8	29	100.0	493	1	YE14_MYCPN	P75372 mycoplasma
9	29	100.0	604	1	YJI3_YEAST	P47030 saccharomyc
10	29	100.0	612	1	ELF1_MOUSE	Q60775 mus musculu
11	29	100.0	619	1	ELF1_HUMAN	P32519 homo sapien
12	29	100.0	919	1	DNL1_HUMAN	P18858 homo sapien
13	29	100.0	985	1	SEC8_DROME	Q9vnh6 drosophila
14	29	100.0	1067	1	BAB2_DROME	Q9w0k4 drosophila
15	29	100.0	1235	1	TRK1_YEAST	P12685 saccharomyc
16	29	100.0	1241	1	TRK1_SACBA	P28569 saccharomyc
17	26	89.7	54	1	RS14_PYRAE	Q8zvw1 pyrobaculum
18	26	89.7	133	1	SECR_MOUSE	Q08535 mus musculu
19	26	89.7	133	1	VEGH_ORFN2	P52584 orf virus (
20	26	89.7	140	1	RT12_DROME	P10735 drosophila
21	26	89.7	191	1	KGUA_SYNEL	Q8dmq7 synechococc
22	26	89.7	193	1	VCO7_ADE04	Q96831 human adeno
23	26	89.7	198	1	VCO7_ADE02	P03266 human adeno
24	26	89.7	199	1	KGUA_ANASP	Q8z0i7 anabaena sp
25	26	89.7	205	1	KGUA_NEIMA	Q9jt96 neisseria m
26	26	89.7	205	1	KGUA_NEIMB	Q9jyb5 neisseria m
27	26	89.7	219	1	Y132_NPVAC	P24730 autographa
28	26	89.7	220	1	PE55_LUCCU	Q95ue8 lucilia cup
29	26	89.7	253	1	NOCT_RAT	Q9et55 rattus norv
30	26	89.7	257	1	TRUA_XYLFA	Q9pdk6 xylella fas
31	26	89.7	257	1	TRUA_XYLFT	Q87ds1 xylella fas
32	26	89.7	304	1	YQOB_CAEEL	Q09300 caenorhabdi
33	26	89.7	350	1	YB04_AQUAE	O67189 aquifex aeo
34	26	89.7	370	1	TRMU_RICCN	Q92il0 rickettsia
35	26	89.7	395	1	RT31_HUMAN	Q92665 homo sapien
36	26	89.7	396	1	TRMU_RHILO	Q98hl0 rhizobium l
37	26	89.7	398	1	TRMU_BRUME	Q8yil6 brucella me
38	26	89.7	398	1	TRMU_BRUSU	Q8cy38 brucella su
39	26	89.7	402	1	VGLD_HSVEA	P24872 equine herp
40	26	89.7	428	1	CCKR_HUMAN	P32238 homo sapien
41	26	89.7	429	1	NOCT_MOUSE	O35710 mus musculu
42	26	89.7	431	1	NOCT_HUMAN	Q9uk39 homo sapien
43	26	89.7	438	1	ENGA_CLOPE	Q8xjk1 clostridium
44	26	89.7	439	1	SRMB_HAEIN	P44701 haemophilus
45	26	89.7	442	1	VGLD_HSVEK	P22484 equine herp
46	26	89.7	452	1	VGLD_HSVEB	P24379 equine herp
47	26	89.7	463	1	ENGA_BIFLO	Q8g6a8 bifidobacte
48	26	89.7	466	1	EM55_HUMAN	Q00013 homo sapien
49	26	89.7	484	1	ENP6_HUMAN	O75354 homo sapien
50	26	89.7	524	1	AMPA_AERPE	Q9y935 aeropyrum p
51	26	89.7	536	1	GAG_MLVDE	P29168 murine leuk
52	26	89.7	552	1	AIRE_MOUSE	Q9z0e3 mus musculu
53	26	89.7	619	1	OM70_NEUCR	P23231 neurospora
54	26	89.7	737	1	POLG_HCVJ5	P27960 hepatitis c
55	26	89.7	737	1	POLG_HCVJ7	P27961 hepatitis c
56	26	89.7	782	1	CS2B_HUMAN	Q9h8e8 homo sapien
57	26	89.7	846	1	IRS1_HCMVA	P09715 human cytom
58	26	89.7	905	1	CTPF_MYCTU	Q10860 mycobacteri
59	26	89.7	928	1	CHS2_EXODE	P30601 exophiala d
60	26	89.7	961	1	LIN2_CAEEL	P54936 caenorhabdi
61	26	89.7	961	1	TSP4_HUMAN	P35443 homo sapien
62	26	89.7	991	1	DHP1_SCHPO	P40848 schizosacch
63	26	89.7	1057	1	POLR_DROME	P16423 drosophila
64	26	89.7	1122	1	VAB1_CAEEL	O61460 caenorhabdi
65	26	89.7	1130	1	ITA6_HUMAN	P23229 homo sapien
66	26	89.7	1160	1	WDR7_HUMAN	Q9y4e6 homo sapien
67	26	89.7	1173	1	DPOL_RCMVM	Q85428 rat cytomeg
68	26	89.7	1446	1	IE18_PRVKA	P33479 pseudorabie
69	26	89.7	1461	1	IE18_PRVIF	P11675 pseudorabie
70	26	89.7	2139	1	CCAC_MOUSE	Q01815 mus musculu
71	26	89.7	2169	1	CCAC_RAT	P22002 rattus norv
72	26	89.7	2171	1	CCAC_RABIT	P15381 oryctolagus
73	26	89.7	2221	1	CCAC_HUMAN	Q13936 homo sapien
74	26	89.7	2353	1	CCAH_HUMAN	O95180 homo sapien
75	26	89.7	3033	1	POLG_HCVJ8	P26661 h genome po
76	26	89.7	5262	1	MLL2_HUMAN	O14686 homo sapien
77	25	86.2	44	1	COXR_RAT	P80433 rattus norv
78	25	86.2	69	1	COXR_BOVIN	P14622 bos taurus
79	25	86.2	111	1	LV2I_HUMAN	P01712 homo sapien
80	25	86.2	114	1	MAUL_METEX	Q49129 methylobact
81	25	86.2	172	1	RUB2_PSEOL	P00272 pseudomonas
82	25	86.2	223	1	GLU2_MAIZE	P04706 zea mays (m
83	25	86.2	228	1	MXI1_HUMAN	P50539 homo sapien
84	25	86.2	228	1	MXI1_MOUSE	P50540 mus musculu
85	25	86.2	228	1	MXI1_RAT	O09015 rattus norv
86	25	86.2	241	1	NGF_HUMAN	P01138 homo sapien
87	25	86.2	262	1	NEF_SIVM1	P05862 simian immu
88	25	86.2	277	1	TRMD_STRCO	O69882 streptomyce
89	25	86.2	278	1	PYRF_CORGL	Q8nq40 corynebacte
90	25	86.2	296	1	YWFM_BACSU	P39649 bacillus su
91	25	86.2	297	1	SGS4_DROME	Q00725 drosophila
92	25	86.2	306	1	FMT_BRUME	Q8ydb3 brucella me
93	25	86.2	311	1	FMT_AGRT5	Q8uid0 agrobacteri
94	25	86.2	311	1	FMT_RHIME	Q92sh5 rhizobium m
95	25	86.2	317	1	FMT_RHILO	Q98d53 rhizobium l
96	25	86.2	318	1	IFR_CICAR	Q00016 cicer ariet
97	25	86.2	332	1	SR4_PHYPO	P11113 physarum po
98	25	86.2	332	1	YIE2_HSVB4	Q02484 bovine herp
99	25	86.2	379	1	GIT7_SCHPO	O59709 schizosacch
100	25	86.2	379	1	PSPB_DICDI	P54704 dictyosteli
101	25	86.2	396	1	TRT_DROME	P19351 drosophila
102	25	86.2	412	1	SU9I_MOUSE	O54864 m histone-l
103	25	86.2	431	1	ACRO_RABIT	P48038 oryctolagus
104	25	86.2	434	1	UL43_HSV11	P10227 herpes simp
105	25	86.2	458	1	ENGA_HELPY	O25505 helicobacte
106	25	86.2	462	1	ENGA_HELPJ	Q9z109 helicobacte

107	25	86.2	469	1	ANTA_GENTR	Q9zwr8 gentiana tr
108	25	86.2	487	1	CPDE_BOVIN	Q01361 bos taurus
109	25	86.2	499	1	CPDF_CANFA	Q29473 canis famil
110	25	86.2	502	1	DCE_LYCES	P54767 lycopersico
111	25	86.2	609	1	YKD6_CAEEL	Q03564 caenorhabdi
112	25	86.2	647	1	TOP3_VIBCH	Q9kqf5 vibrio chol
113	25	86.2	649	1	TOP3_SALTI	Q8z6f5 salmonella
114	25	86.2	649	1	TOP3_SALTY	P40687 salmonella
115	25	86.2	653	1	TOP3_ECOLI	P14294 escherichia
116	25	86.2	672	1	NGR1_YEAST	P32831 saccharomyc
117	25	86.2	679	1	ENV2_MOUSE	P11370 mus musculu
118	25	86.2	908	1	SRCA_RABIT	P13666 oryctolagus
119	25	86.2	988	1	ST23_YEAST	Q06010 saccharomyc
120	25	86.2	1035	1	CC68_YEAST	P32558 saccharomyc
121	25	86.2	1074	1	EMI2_MOUSE	Q8k482 mus musculu
122	25	86.2	1080	1	NRK1_YEAST	P38692 saccharomyc
123	25	86.2	1152	1	ITAM_HUMAN	P11215 homo sapien
124	25	86.2	1175	1	DSRA_RAT	P55266 rattus norv
125	25	86.2	1210	1	AF4_HUMAN	P51825 homo sapien
126	25	86.2	1276	1	BXD_CLOBO	P19321 clostridium
127	25	86.2	1411	1	TCOF_HUMAN	Q13428 homo sapien
128	25	86.2	1484	1	CES2_HUMAN	Q9bxf3 homo sapien
129	25	86.2	1556	1	Y934_HUMAN	Q9y2e4 homo sapien
130	25	86.2	2146	1	INSR_DROME	P09208 drosophila
131	25	86.2	2472	1	NCR2_MOUSE	Q9wu42 mus musculu
132	25	86.2	5596	1	MDN1_HUMAN	Q9nu22 homo sapien
133	24	82.8	57	1	Y23A_MYCLE	Q9cd20 mycobacteri
134	24	82.8	59	1	MAMB_DENJA	P28375 dendroaspis
135	24	82.8	60	1	RZOD_ECOLI	P58041 escherichia
136	24	82.8	61	1	TX51_DENJA	P01413 dendroaspis
137	24	82.8	62	1	ITR1_ASCSU	P19398 ascaris suu
138	24	82.8	63	1	YAIA_ECOLI	P08366 escherichia
139	24	82.8	67	1	SLYX_RHILO	Q98gw1 rhizobium l
140	24	82.8	68	1	SLYX_BRUME	Q8yde6 brucella me
141	24	82.8	68	1	SLYX_BRUSU	Q8fuz1 brucella su
142	24	82.8	73	1	UL12_HCMVA	P16777 human cytom
143	24	82.8	80	1	SSS2_SCYCA	P11020 scyliorhinu
144	24	82.8	91	1	RL31_PYRAE	Q8ztu2 pyrobaculum
145	24	82.8	93	1	N12A_MEDSA	Q40361 medicago sa
146	24	82.8	96	1	YBCO_ECOLI	Q47271 escherichia
147	24	82.8	103	1	NO12_MEDTR	P30365 medicago tr
148	24	82.8	105	1	THGF_TOBAC	P32026 nicotiana t
149	24	82.8	106	1	THG1_NICPA	O24115 nicotiana p
150	24	82.8	107	1	RBX1_SCHPO	O13959 schizosacch

ALIGNMENTS

```
RESULT 1
D127_HUMAN
ID   D127_HUMAN     STANDARD;      PRT;    99 AA.
AC   Q9H1M4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Beta-defensin 127 precursor (Beta-defensin 27) (DEFB-27).
GN   DEFB127 OR DEFB27 OR C20ORF73.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
RN   [2]
RP   SEQUENCE OF 19-77 FROM N.A., AND IDENTIFICATION.
RC   TISSUE=B-cell, Fetal lung, and Testis;
RX   MEDLINE=21843921; PubMed=11854508;
RA   Schutte B.C., Mitros J.P., Bartlett J.A., Walters J.D., Jia H.P.,
RA   Welsh M.J., Casavant T.L., McCray P.B. Jr.;
RT   "Discovery of five conserved beta-defensin gene clusters using a
RT   computational search strategy.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:2129-2133(2002).
CC   -!- FUNCTION: Has antibacterial activity (Potential).
CC   -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC   -!- SIMILARITY: Belongs to the beta-defensin family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AL360078; CAC17685.1; -.
DR   EMBL; AY122479; AAM93920.1; -.
DR   Genew; HGNC:16206; DEFB127.
DR   GO; GO:0003797; F:antibacterial peptide activity; TAS.
DR   GO; GO:0045087; P:innate immune response; TAS.
KW   Antibiotic; Signal.
FT   SIGNAL        1     18      POTENTIAL.
FT   CHAIN        19     63      BETA-DEFENSIN 127.
FT   PROPEP       66     99      POTENTIAL.
FT   DISULFID     24     53      BY SIMILARITY.
FT   DISULFID     33     47      BY SIMILARITY.
FT   DISULFID     37     54      BY SIMILARITY.
FT   CONFLICT     31     31      G -> R (IN REF. 2).
FT   CONFLICT     71     77      RPKPATL -> SPKQHLH (IN REF. 2).
SQ   SEQUENCE   99 AA;  11342 MW;  C9CBDB27D17AF380 CRC64;

  Query Match             100.0%;  Score 29;  DB 1;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         67 TKPPR 71


RESULT 2
COAE_SYNY3
ID   COAE_SYNY3     STANDARD;      PRT;   201 AA.
AC   Q55515;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Dephospho-CoA kinase (EC 2.7.1.24) (Dephosphocoenzyme A kinase).
```

```
GN   COAE OR SLR0553.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
CC   -!- FUNCTION: Catalyzes the phosphorylation of the 3'-hydroxyl group
CC       of dephosphocoenzyme A to form coenzyme A (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + dephospho-CoA = ADP + CoA.
CC   -!- PATHWAY: Coenzyme A (CoA) biosynthesis; fifth (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the coaE family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D64006; BAA10873.1; -.
DR   PIR; S76026; S76026.
DR   HAMAP; MF_00376; -; 1.
DR   InterPro; IPR001977; Depp_CoAkinase.
DR   Pfam; PF01121; CoaE; 1.
DR   ProDom; PD003329; Depp_CoAkinase; 1.
DR   TIGRFAMs; TIGR00152; TIGR00152; 1.
DR   PROSITE; PS01294; COAE; 1.
KW   Transferase; Kinase; ATP-binding; Coenzyme A biosynthesis;
KW   Complete proteome.
FT   NP_BIND      15     22       ATP (POTENTIAL).
SQ   SEQUENCE   201 AA;  22520 MW;  34C3E142337F519E CRC64;

  Query Match           100.0%;  Score 29;  DB 1;  Length 201;
  Best Local Similarity  100.0%;  Pred. No. 45;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          3 TKPPR 7


RESULT 3
FA3C_HUMAN
ID   FA3C_HUMAN     STANDARD;      PRT;   227 AA.
AC   Q92520;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Protein FAM3C precursor (Protein GS3786).
GN   FAM3C.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bone;
RA   Ohno I., Hashimoto J., Takaoka K., Ochi T., Okubo K., Matsubara K.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Ryan E., Bauer C., Tucci S., Spalding L.;
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Duodenum;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=22150867; PubMed=12160727;
RA   Zhu Y., Xu G., Patel A., McLaughlin M.M., Silverman C., Knecht K.A.,
RA   Sweitzer S., Li X., McDonnell P., Mirabile R., Zimmerman D., Boyce R.,
RA   Tierney L.A., Hu E., Livi G.P., Wolf B.A., Abdel-Meguid S.S.,
RA   Rose G.D., Aurora R., Hensley P., Briggs M., Young P.R.;
RT   "Cloning, expression, and initial characterization of a novel
RT   cytokine-like gene family.";
RL   Genomics 80:144-150(2002).
CC   -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC   -!- TISSUE SPECIFICITY: Ubiquitous.
CC   -!- SIMILARITY: Belongs to the FAM3 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D87120; BAA13251.1; -.
DR   EMBL; AC006364; AAQ96872.1; -.
DR   EMBL; BC046932; AAH46932.1; -.
DR   Genew; HGNC:18664; FAM3C.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0005125; F:cytokine activity; NAS.
KW   Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    227       PROTEIN FAM3C.
FT   DISULFID     58    221       POTENTIAL.
FT   DISULFID     64     86       POTENTIAL.
SQ   SEQUENCE   227 AA;  24680 MW;  6DC94B259052647F CRC64;

  Query Match           100.0%;  Score 29;  DB 1;  Length 227;
  Best Local Similarity  100.0%;  Pred. No. 50;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         51 TKPPR 55


RESULT 4
FA3C_MOUSE
ID   FA3C_MOUSE     STANDARD;      PRT;   227 AA.
AC   Q91VU0; Q9CTB4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
```

```
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Protein FAM3C precursor.
GN   FAM3C OR D6WSU176E.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE OF 127-227 FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC   -!- SIMILARITY: Belongs to the FAM3 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; BC009086; AAH09086.1; -.
DR   EMBL; AK004059; BAB23146.1; -.
DR   MGD; MGI:107892; D6Wsu176e.
DR   GO; GO:0005576; C:extracellular; ISS.
DR   GO; GO:0005125; F:cytokine activity; ISS.
KW   Signal.
FT   SIGNAL       1     24       POTENTIAL.
FT   CHAIN        25    227      PROTEIN FAM3C.
FT   DISULFID     58    221      POTENTIAL.
FT   DISULFID     64    86       POTENTIAL.
SQ   SEQUENCE   227 AA;  24752 MW;  9195280B92838CF4 CRC64;

  Query Match             100.0%;  Score 29;  DB 1;  Length 227;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         51 TKPPR 55


RESULT 5
RPC_BP163
ID   RPC_BP163      STANDARD;      PRT;   262 AA.
AC   P15238; Q9MCD2;
DT   01-APR-1990 (Rel. 14, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Repressor protein C.
GN   C.
OS   Bacteriophage 16-3.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales.
OX   NCBI_TaxID=10704;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-3.
RA   Dallmann G., Papp P., Orosz L.;
RT   "Related repressor specificity of unrelated phages.";
RL   Nature 330:398-401(1987).
RN   [2]
RP   REVISIONS.
RX   MEDLINE=99328962; PubMed=10400574;
RA   Semsey S., Papp I., Buzas Z., Patthy A., Orosz L., Papp P.P.;
RT   "Identification of site-specific recombination genes int and xis of
RT   the Rhizobium temperate phage 16-3.";
RL   J. Bacteriol. 181:4185-4192(1999).
CC   -!- SIMILARITY: Contains 1 HTH cro/C1-type DNA-binding domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ131679; CAB54833.1; -.
DR   PIR; S01612; RPBP16.
DR   InterPro; IPR001387; HTH_3.
DR   Pfam; PF01381; HTH_3; 1.
DR   SMART; SM00530; HTH_XRE; 1.
DR   PROSITE; PS50943; HTH_CROC1; 1.
KW   Early protein; Transcription regulation; Repressor; DNA-binding.
FT   INIT_MET     0      0
FT   DOMAIN       14     67       HTH CRO/C1-TYPE.
FT   DNA_BIND     25     44       H-T-H MOTIF (PROBABLE).
SQ   SEQUENCE   262 AA;  29522 MW;  56DE1FF97AC9C010 CRC64;

  Query Match             100.0%;  Score 29;  DB 1;  Length 262;
  Best Local Similarity   100.0%;  Pred. No. 58;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        192 TKPPR 196


RESULT 6
ENGA_STRCO
ID   ENGA_STRCO     STANDARD;      PRT;   465 AA.
AC   Q9EWW8;
```

```
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   GTP-binding protein engA.
GN   ENGA OR SCO1758 OR 2SCI34.11C.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
CC   -!- FUNCTION: GTPase of unknown physiological role.
CC   -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
CC       proteins. EngA subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL939110; CAC12931.1; ALT_INIT.
DR   HAMAP; MF_00195; -; 1.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR005289; GTP-bindding_dom.
DR   InterPro; IPR006073; GTP1_OBG.
DR   InterPro; IPR002917; MMR_HSR1.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF01926; MMR_HSR1; 1.
DR   PRINTS; PR00326; GTP1OBG.
DR   SMART; SM00382; AAA; 2.
DR   TIGRFAMs; TIGR00650; MG442; 2.
DR   TIGRFAMs; TIGR00231; small_GTP; 2.
KW   GTP-binding; Repeat; Complete proteome.
FT   NP_BIND      33     40       GTP 1 (POTENTIAL).
FT   NP_BIND      80     84       GTP 1 (POTENTIAL).
FT   NP_BIND     142    145       GTP 1 (POTENTIAL).
FT   NP_BIND     208    215       GTP 2 (POTENTIAL).
FT   NP_BIND     255    259       GTP 2 (POTENTIAL).
FT   NP_BIND     320    323       GTP 2 (POTENTIAL).
SQ   SEQUENCE   465 AA;  50344 MW;  9E053E7C00C859F8 CRC64;

  Query Match           100.0%;  Score 29;  DB 1;  Length 465;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        414 TKPPR 418


RESULT 7
ENGA_STRAW
ID   ENGA_STRAW     STANDARD;      PRT;   491 AA.
AC   Q828Y7;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   GTP-binding protein engA.
GN   ENGA OR SAV6524.
OS   Streptomyces avermitilis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=33903;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=21477403; PubMed=11572948;
RA   Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA   Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=22608306; PubMed=12692562;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industrial
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
CC   -!- FUNCTION: GTPase of unknown physiological role.
CC   -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
CC       proteins. EngA subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP005047; BAC74235.1; -.
DR   HAMAP; MF_00195; -; 1.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR005289; GTP-bindding_dom.
DR   InterPro; IPR006073; GTP1_OBG.
DR   InterPro; IPR002917; MMR_HSR1.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF01926; MMR_HSR1; 1.
DR   PRINTS; PR00326; GTP1OBG.
DR   SMART; SM00382; AAA; 2.
DR   TIGRFAMs; TIGR00650; MG442; 2.
DR   TIGRFAMs; TIGR00231; small_GTP; 2.
KW   GTP-binding; Repeat; Complete proteome.
FT   NP_BIND      59     66       GTP 1 (POTENTIAL).
FT   NP_BIND     106    110       GTP 1 (POTENTIAL).
FT   NP_BIND     168    171       GTP 1 (POTENTIAL).
FT   NP_BIND     234    241       GTP 2 (POTENTIAL).
FT   NP_BIND     281    285       GTP 2 (POTENTIAL).
FT   NP_BIND     346    349       GTP 2 (POTENTIAL).
SQ   SEQUENCE   491 AA;  53353 MW;  7BADD8E0449D14AC CRC64;

  Query Match           100.0%;  Score 29;  DB 1;  Length 491;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        440 TKPPR 444


RESULT 8
YE14_MYCPN
ID   YE14_MYCPN     STANDARD;      PRT;   493 AA.
AC   P75372;
```

```
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Very hypothetical mgpC-like protein MPN414 (A05_orf493).
GN   MPN414 OR MP426.
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2104;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE=97105885; PubMed=8948633;
RA   Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA   Herrmann R.;
RT   "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT   pneumoniae.";
RL   Nucleic Acids Res. 24:4420-4449(1996).
CC   -!- SIMILARITY: BELONGS TO THE MGPC FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000041; AAB96074.1; -.
DR   PIR; S73752; S73752.
DR   InterPro; IPR007885; MgpC.
DR   Pfam; PF05220; MgpC; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   493 AA;  52836 MW;  535607770E515304 CRC64;

  Query Match             100.0%;  Score 29;  DB 1;  Length 493;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        462 TKPPR 466


RESULT 9
YJI3_YEAST
ID   YJI3_YEAST     STANDARD;      PRT;   604 AA.
AC   P47030;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Hypothetical 68.8 kDa protein in EXO70-ARP4 intergenic region.
GN   YJL083W OR J1002.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=96093911; PubMed=7483841;
RA   Miosga T., Schaaff-Gerstenschlaeger I., Chalwatzis N., Baur A.,
RA   Boles E., Fournier C., Schmitt S., Velten C., Wilhelm N.,
RA   Zimmermann F.K.;
RT   "Sequence analysis of a 33.1 kb fragment from the left arm of
RT   Saccharomyces cerevisiae chromosome X, including putative proteins
RT   with leucine zippers, a fungal Zn(II)2-Cys6 binuclear cluster domain
RT   and a putative alpha 2-SCB-alpha 2 binding site.";
RL   Yeast 11:681-689(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RA   Sor F.J.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO YEAST YKR019C.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z49358; CAA89376.1; -.
DR   EMBL; X83502; CAA58487.1; -.
DR   EMBL; X88851; CAA61318.1; -.
DR   PIR; S56027; S56027.
DR   GermOnline; 141697; -.
DR   SGD; S0003619; YJL083W.
DR   InterPro; IPR000261; EPS15_homology.
DR   SMART; SM00027; EH; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   604 AA;  68768 MW;  E17C6C71D9FBFBD6 CRC64;

  Query Match             100.0%;  Score 29;  DB 1;  Length 604;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        151 TKPPR 155


RESULT 10
ELF1_MOUSE
ID   ELF1_MOUSE     STANDARD;      PRT;   612 AA.
AC   Q60775;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   ETS-related transcription factor Elf-1 (E74-like factor 1).
GN   ELF1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=96257231; PubMed=8666284;
RA   Davis J.N., Roussel M.F.;
RT   "Cloning and expression of the murine Elf-1 cDNA.";
RL   Gene 171:265-269(1996).
CC   -!- FUNCTION: TRANSCRIPTION FACTOR THAT APPEARS TO BE REQUIRED FOR THE
CC       T-CELL-RECEPTOR-MEDIATED TRANS ACTIVATION OF HIV-2 GENE
CC       EXPRESSION. BINDS SPECIFICALLY TO TWO PURINE-RICH MOTIFS IN THE
CC       HIV-2 ENHANCER. ELF-1 BINDS TO THE UNDERPHOSPHORYLATED FORM OF RB.
CC       MAY INTERACT WITH OTHER TRANSCRIPTION FACTORS IN ORDER TO REGULATE
CC       SPECIFIC GENES (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY FOUND IN HEMATOPOIETIC CELLS.
CC       DETECTED IN OTHER CELL TYPES SUCH AS FIBROBLASTS.
CC   -!- SIMILARITY: Belongs to the ETS family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U19617; AAB17097.1; -.
DR   HSSP; P28324; 1BC8.
DR   TRANSFAC; T05012; -.
DR   MGD; MGI:107180; Elf1.
```

```
DR   InterPro; IPR000418; Ets.
DR   InterPro; IPR002341; HSF_ETS.
DR   Pfam; PF00178; Ets; 1.
DR   PRINTS; PR00454; ETSDOMAIN.
DR   SMART; SM00413; ETS; 1.
DR   PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR   PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR   PROSITE; PS50061; ETS_DOMAIN_3; 1.
KW   Nuclear protein; Transcription regulation; Activator; DNA-binding.
FT   DOMAIN       75     80       POLY-ASP.
FT   DNA_BIND     208    290      ETS-DOMAIN.
SQ   SEQUENCE   612 AA;  66221 MW;  442F4C85142B31F0 CRC64;

  Query Match             100.0%;  Score 29;  DB 1;  Length 612;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        180 TKPPR 184


RESULT 11
ELF1_HUMAN
ID   ELF1_HUMAN     STANDARD;      PRT;   619 AA.
AC   P32519; Q9UDE1;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   ETS-related transcription factor Elf-1 (E74-like factor 1).
GN   ELF1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92407982; PubMed=1527846;
RA   Leiden J.M., Wang C.Y., Petryniak B., Markovitz D.M., Nabel G.J.,
RA   Thompson C.B.;
RT   "A novel Ets-related transcription factor, Elf-1, binds to human
RT   immunodeficiency virus type 2 regulatory elements that are required
RT   for inducible trans activation in T cells.";
RL   J. Virol. 66:5890-5897(1992).
RN   [2]
RP   SEQUENCE OF 204-289 FROM N.A.
RX   MEDLINE=92186836; PubMed=1545787;
RA   Thompson C.B., Wang C.Y., Ho I.C., Bohjanen P.R., Petryniak B.,
RA   June C.H., Miesfeldt S., Zhang L., Nabel G.J., Karpinski B.;
RT   "Cis-acting sequences required for inducible interleukin-2 enhancer
RT   function bind a novel Ets-related protein, Elf-1.";
RL   Mol. Cell. Biol. 12:1043-1053(1992).
RN   [3]
RP   BINDING TO RB.
RX   MEDLINE=93262492; PubMed=8493578;
RA   Wang C.Y., Petryniak B., Thompson C.B., Kaelin W.G., Leiden J.M.;
RT   "Regulation of the Ets-related transcription factor Elf-1 by binding
RT   to the retinoblastoma protein.";
RL   Science 260:1330-1335(1993).
CC   -!- FUNCTION: TRANSCRIPTION FACTOR THAT APPEARS TO BE REQUIRED FOR THE
CC       T-CELL-RECEPTOR-MEDIATED TRANS ACTIVATION OF HIV-2 GENE
CC       EXPRESSION. BINDS SPECIFICALLY TO TWO PURINE-RICH MOTIFS IN THE
CC       HIV-2 ENHANCER. ELF-1 BINDS TO THE UNDERPHOSPHORYLATED FORM OF RB.
CC       MAY INTERACT WITH OTHER TRANSCRIPTION FACTORS IN ORDER TO REGULATE
CC       SPECIFIC GENES.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the ETS family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M82882; -; NOT_ANNOTATED_CDS.
DR   PIR; A43361; A43361.
DR   HSSP; P28324; 1BC8.
DR   TRANSFAC; T01113; -.
DR   Genew; HGNC:3316; ELF1.
DR   MIM; 189973; -.
DR   GO; GO:0005634; C:nucleus; NAS.
DR   GO; GO:0016563; F:transcriptional activator activity; NAS.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR   InterPro; IPR000418; Ets.
DR   InterPro; IPR002341; HSF_ETS.
DR   Pfam; PF00178; Ets; 1.
DR   PRINTS; PR00454; ETSDOMAIN.
DR   SMART; SM00413; ETS; 1.
DR   PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR   PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR   PROSITE; PS50061; ETS_DOMAIN_3; 1.
KW   Nuclear protein; Transcription regulation; Activator; DNA-binding.
FT   DOMAIN       75     80       POLY-ASP.
FT   DNA_BIND     208    290      ETS-DOMAIN.
FT   CONFLICT     283    283      Q -> G (IN REF. 2).
SQ   SEQUENCE   619 AA;  67455 MW;  AB0B41B2964A66EF CRC64;

  Query Match             100.0%;  Score 29;  DB 1;  Length 619;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        180 TKPPR 184


RESULT 12
DNL1_HUMAN
ID   DNL1_HUMAN     STANDARD;      PRT;   919 AA.
AC   P18858;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   DNA ligase I (EC 6.5.1.1) (Polydeoxyribonucleotide synthase [ATP]).
GN   LIG1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=T lymphoblast;
RX   MEDLINE=90370849; PubMed=2204063;
RA   Barnes D.E., Johnston L.H., Kodama K.I., Tomkinson A.E.,
RA   Lasko D.D., Lindahl T.;
RT   "Human DNA ligase I cDNA: cloning and functional expression in
RT   Saccharomyces cerevisiae.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:6679-6683(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,
RA   Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA   Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 716-753 FROM N.A.
RX   MEDLINE=91352039; PubMed=1881902;
RA   Petrini J.H.J., Huwiler K.G., Weaver D.T.;
RT   "A wild-type DNA ligase I gene is expressed in Bloom's syndrome
RT   cells.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:7615-7619(1991).
RN   [4]
RP   VARIANTS LYS-566 AND TRP-771.
RX   MEDLINE=92257590; PubMed=1581963;
```

```
RA   Barnes D.E., Tomkinson A.E., Lehmann A.R., Webster A.D.B.,
RA   Lindahl T.;
RT   "Mutations in the DNA ligase I gene of an individual with
RT   immunodeficiencies and cellular hypersensitivity to DNA-damaging
RT   agents.";
RL   Cell 69:495-503(1992).
CC   -!- FUNCTION: This protein seals, during DNA replication, DNA
CC       recombination and DNA repair, nicks in double-stranded DNA.
CC   -!- CATALYTIC ACTIVITY: ATP + (deoxyribonucleotide)(N) +
CC       (deoxyribonucleotide)(M) = AMP + diphosphate +
CC       (deoxyribonucleotide)(N+M).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- PTM: Phosphorylated in vivo.
CC   -!- DISEASE: Defects in LIG1 seem to cause immunodeficiencies and
CC       cellular hypersensitivity to DNA-damaging agents.
CC   -!- SIMILARITY: Belongs to the ATP-dependent DNA ligase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M36067; AAA59518.1; -.
DR   EMBL; AF527418; AAM77697.1; -.
DR   PIR; A36048; A41275.
DR   Genew; HGNC:6598; LIG1.
DR   MIM; 126391; -.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0003677; F:DNA binding; TAS.
DR   GO; GO:0003909; F:DNA ligase activity; TAS.
DR   GO; GO:0006281; P:DNA repair; TAS.
DR   GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR   InterPro; IPR000977; DNA_ligase.
DR   Pfam; PF01068; DNA_ligase; 1.
DR   Pfam; PF04679; DNA_ligase_A_C; 1.
DR   Pfam; PF04675; DNA_ligase_A_N; 1.
DR   TIGRFAMs; TIGR00574; dnl1; 1.
DR   PROSITE; PS00697; DNA_LIGASE_A1; 1.
DR   PROSITE; PS00333; DNA_LIGASE_A2; 1.
DR   PROSITE; PS50160; DNA_LIGASE_A3; 1.
KW   DNA repair; DNA replication; DNA recombination; Cell division; Ligase;
KW   ATP-binding; Nuclear protein; Phosphorylation; Polymorphism;
KW   Disease mutation.
FT   BINDING     568    568       AMP (BY SIMILARITY).
FT   VARIANT     249    249       G -> E (in dbSNP:3730911).
FT                                /FTId=VAR_016766.
FT   VARIANT     267    267       N -> S (in dbSNP:3730933).
FT                                /FTId=VAR_016767.
FT   VARIANT     409    409       R -> H (in dbSNP:4987068).
FT                                /FTId=VAR_016768.
FT   VARIANT     480    480       M -> V (in dbSNP:3730980).
FT                                /FTId=VAR_016769.
FT   VARIANT     566    566       E -> K (in LIG1 deficiency).
FT                                /FTId=VAR_002262.
FT   VARIANT     614    614       T -> I (in dbSNP:3731003).
FT                                /FTId=VAR_016770.
FT   VARIANT     771    771       R -> W (in LIG1 deficiency).
FT                                /FTId=VAR_002263.
SQ   SEQUENCE   919 AA;  101735 MW;  B2854DAE38A8D4AD CRC64;

  Query Match             100.0%;  Score 29;  DB 1;  Length 919;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        218 TKPPR 222


RESULT 13
```

```
SEC8_DROME
ID   SEC8_DROME     STANDARD;      PRT;   985 AA.
AC   Q9VNH6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Probable exocyst complex component Sec8.
GN   CG2095.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   REVISIONS.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley; TISSUE=Embryo;
RX   MEDLINE=22426066; PubMed=12537569;
RA   Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
```

```
RA   George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA   Rubin G.M., Celniker S.E.;
RT   "A Drosophila full-length cDNA resource.";
RL   Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC   -!- FUNCTION: Component of the exocyst complex involved in the docking
CC       of exocystic vesicles with fusions site on the plasma membrane (By
CC       similarity).
CC   -!- SUBUNIT: The exocyst complex is composed of SEC3, SEC5, SEC6,
CC       SEC8, SEC10, SEC15, EXO70 and EXO84 (By similarity).
CC   -!- SIMILARITY: Belongs to the SEC8 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE003601; AAF51959.3; -.
DR   EMBL; AY119660; AAM50314.1; -.
DR   FlyBase; FBgn0037373; CG2095.
DR   InterPro; IPR004172; L27.
DR   InterPro; IPR007191; Sec8_exocyst.
DR   Pfam; PF02828; L27; 1.
DR   Pfam; PF04048; Sec8_exocyst; 1.
KW   Exocytosis; Transport; Protein transport; Coiled coil.
FT   DOMAIN       36     70       COILED COIL (POTENTIAL).
FT   CONFLICT    118    118       M -> T (IN REF. 3).
FT   CONFLICT    209    209       E -> D (IN REF. 3).
FT   CONFLICT    333    333       S -> T (IN REF. 3).
FT   CONFLICT    668    668       Y -> S (IN REF. 3).
SQ   SEQUENCE   985 AA;  111665 MW;  6BFB0D9C539FEBE7 CRC64;

  Query Match             100.0%;  Score 29;  DB 1;  Length 985;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          7 TKPPR 11


RESULT 14
BAB2_DROME
ID   BAB2_DROME     STANDARD;      PRT;  1067 AA.
AC   Q9W0K4; Q24001; Q9U1H3;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Bric-a-brac protein 2.
GN   BAB2 OR BTBII OR CG9102/CG13911.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
RP   SPECIFICITY.
RC   TISSUE=Embryo, and Ovary;
RX   MEDLINE=21969340; PubMed=11973274;
RA   Couderc J.L.G., Godt D., Zollman S., Chen J., Li M., Tiong S.,
RA   Cramton S.E., Sahut-Barnola I., Laski F.A.;
RT   "The bric a brac locus consists of two paralogous genes encoding
RT   BTB/POZ domain proteins and acts as a homeotic and morphogenetic
RT   regulator of imaginal development in Drosophila.";
RL   Development 129:2419-2433(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   REVISIONS.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [4]
RP   SEQUENCE OF 196-310 FROM N.A.
RX   MEDLINE=95024186; PubMed=7938017;
RA   Zollman S., Godt D., Prive G.G., Couderc J.L., Laski F.A.;
RT   "The BTB domain, found primarily in zinc finger proteins, defines an
RT   evolutionarily conserved family that includes several developmentally
RT   regulated genes in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:10717-10721(1994).
CC   -!- FUNCTION: Probably acts as a transcriptional regulator. Required
CC       for the specification of the tarsal segment. Also involved in
CC       antenna development.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: Leg imaginal disk at the central region of the
CC       tarsus and in eye antenna disk at the basal cylinder.
CC   -!- MISCELLANEOUS: 'bric-a-brac' means 'jumble' in French (refering to
CC       the mutant ovary phenotype).
CC   -!- SIMILARITY: Contains 1 A.T hook DNA-binding repeat.
CC   -!- SIMILARITY: Contains 1 BTB/POZ domain.
CC   -!- SIMILARITY: Contains 1 helix-turn-helix Psq-type domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ252173; CAB64388.1; -.
DR   EMBL; AE003470; AAF47442.2; -.
DR   EMBL; U14399; AAA50834.1; -.
DR   FlyBase; FBgn0025525; bab2.
DR   InterPro; IPR000637; AT_hook.
DR   InterPro; IPR000210; BTB_POZ.
DR   Pfam; PF00651; BTB; 1.
DR   SMART; SM00225; BTB; 1.
DR   PROSITE; PS50097; BTB; 1.
KW   Nuclear protein; DNA-binding; Transcription regulation.
FT   DOMAIN      223    288       BTB.
FT   DNA_BIND    645    690       H-T-H MOTIF PSQ-TYPE.
FT   DNA_BIND    697    708       A.T HOOK.
FT   CONFLICT    858    858       A -> R (IN REF. 1).
SQ   SEQUENCE   1067 AA;  114661 MW;  7DBFC7681D507FC0 CRC64;

  Query Match             100.0%;  Score 29;  DB 1;  Length 1067;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        181 TKPPR 185


RESULT 15
TRK1_YEAST
ID   TRK1_YEAST     STANDARD;      PRT;  1235 AA.
AC   P12685;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Potassium transport protein, high-affinity.
GN   TRK1 OR YJL129C OR J0693.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88302204; PubMed=3043197;
RA   Gaber R.F., Styles C.A., Fink G.R.;
RT   "TRK1 encodes a plasma membrane protein required for high-affinity
RT   potassium transport in Saccharomyces cerevisiae.";
RL   Mol. Cell. Biol. 8:2848-2859(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / FY1679;
RX   MEDLINE=97103775; PubMed=8948101;
RA   Cziepluch C., Kordes E., Pujol A., Jauniaux J.-C.;
RT   "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
RT   reveals 19 open reading frames including URA2 (5' end), TRK1, PBS2,
RT   SPT10, GCD14, RPE1, PHO86, NCA3, ASF1, CCT7, GZF3, two tRNA genes,
RT   three remnant delta elements and a Ty4 transposon.";
RL   Yeast 12:1471-1474(1996).
CC   -!- FUNCTION: This protein is required for high-affinity potassium
CC       transport.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the trkH potassium transport family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21328; AAA34728.1; -.
DR   EMBL; Z49404; CAA89424.1; -.
DR   PIR; S05849; PWBYH.
DR   GermOnline; 141741; -.
DR   SGD; S0003665; TRK1.
DR   GO; GO:0015079; F:potassium ion transporter activity; IDA.
DR   GO; GO:0030007; P:potassium ion homeostasis; IDA.
DR   InterPro; IPR003445; Cat_transpt.
DR   InterPro; IPR004773; Ktransp_euk.
DR   Pfam; PF02386; TrkH; 1.
DR   PIRSF; PIRSF002450; K+_transpter_TRK; 1.
DR   TIGRFAMs; TIGR00934; 2a38euk; 1.
KW   Transport; Transmembrane; Potassium transport; Glycoprotein.
FT   TRANSMEM     49     70       POTENTIAL.
FT   TRANSMEM     78     98       POTENTIAL.
FT   TRANSMEM    107    127       POTENTIAL.
FT   TRANSMEM    778    800       POTENTIAL.
FT   TRANSMEM    813    834       POTENTIAL.
FT   TRANSMEM    838    858       POTENTIAL.
FT   TRANSMEM    862    882       POTENTIAL.
FT   TRANSMEM    898    918       POTENTIAL.
FT   TRANSMEM    923    943       POTENTIAL.
FT   TRANSMEM    971    991       POTENTIAL.
FT   TRANSMEM   1078   1098       POTENTIAL.
FT   TRANSMEM   1111   1131       POTENTIAL.
FT   CARBOHYD    100    100       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    169    169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    222    222       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    227    227       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    251    251       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    369    369       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    383    383       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    497    497       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    501    501       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    532    532       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    580    580       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    677    677       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    919    919       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1030   1030       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1135   1135       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1235 AA;  141072 MW;  BCE9FDBA0BA0982B CRC64;

  Query Match             100.0%;  Score 29;  DB 1;  Length 1235;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        438 TKPPR 442


RESULT 16
TRK1_SACBA
ID   TRK1_SACBA     STANDARD;      PRT;  1241 AA.
AC   P28569;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Potassium transport protein, high-affinity.
GN   TRK1.
OS   Saccharomyces bayanus (Yeast) (Saccharomyces uvarum).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4931;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1668;
RX   MEDLINE=91216443; PubMed=2022322;
RA   Anderson J.A., Best L.A., Gaber R.F.;
RT   "Structural and functional conservation between the high-affinity K+
RT   transporters of Saccharomyces uvarum and Saccharomyces cerevisiae.";
RL   Gene 99:39-46(1991).
CC   -!- FUNCTION: This protein is required for high-affinity potassium
```

```
CC       transport.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the trkH potassium transport family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57508; AAA34661.1; -.
DR   InterPro; IPR003445; Cat_transpt.
DR   InterPro; IPR004773; Ktransp_euk.
DR   Pfam; PF02386; TrkH; 1.
DR   PIRSF; PIRSF002450; K+_transpter_TRK; 1.
DR   TIGRFAMs; TIGR00934; 2a38euk; 1.
KW   Transport; Transmembrane; Potassium transport; Glycoprotein.
FT   TRANSMEM     49     70       POTENTIAL.
FT   TRANSMEM     78     98       POTENTIAL.
FT   TRANSMEM    107    127       POTENTIAL.
FT   TRANSMEM    784    806       POTENTIAL.
FT   TRANSMEM    819    840       POTENTIAL.
FT   TRANSMEM    844    864       POTENTIAL.
FT   TRANSMEM    868    888       POTENTIAL.
FT   TRANSMEM    904    924       POTENTIAL.
FT   TRANSMEM    929    949       POTENTIAL.
FT   TRANSMEM    977    997       POTENTIAL.
FT   TRANSMEM   1084   1104       POTENTIAL.
FT   TRANSMEM   1117   1137       POTENTIAL.
FT   CARBOHYD    100    100       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    223    223       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    227    227       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    233    233       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    257    257       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    274    274       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    353    353       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    364    364       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    389    389       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    442    442       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    505    505       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    538    538       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    584    584       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    660    660       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    681    681       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    691    691       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    741    741       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    925    925       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1141   1141       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1241 AA;  141225 MW;  F5149E47EC00BCD2 CRC64;
```

```
  Query Match             100.0%;  Score 29;  DB 1;  Length 1241;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        444 TKPPR 448
```

```
RESULT 17
RS14_PYRAE
ID   RS14_PYRAE     STANDARD;      PRT;    54 AA.
AC   Q8ZVW1; P58732;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S14P.
GN   RPS14P OR PAE2097.
OS   Pyrobaculum aerophilum.
OC   Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC   Thermoproteaceae; Pyrobaculum.
OX   NCBI_TaxID=13773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX   MEDLINE=21664397; PubMed=11792869;
RA   Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA   Miller J.H.;
RT   "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT   aerophilum.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
CC   -!- SIMILARITY: Belongs to the S14P family of ribosomal proteins.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE009857; AAL63943.1; -.
DR   InterPro; IPR001209; Ribosomal_S14.
DR   Pfam; PF00253; Ribosomal_S14; 1.
DR   PROSITE; PS00527; RIBOSOMAL_S14; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   54 AA;  6444 MW;  153B7BB801EDD963 CRC64;
```

```
  Query Match             89.7%;  Score 26;  DB 1;  Length 54;
  Best Local Similarity   80.0%;  Pred. No. 47;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db          4 TKPPK 8
```

```
RESULT 18
SECR_MOUSE
ID   SECR_MOUSE     STANDARD;      PRT;   133 AA.
AC   Q08535;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Secretin precursor.
GN   SCT.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94234995; PubMed=8179583;
RA   Lan M.S., Kajiyama W., Donadel G., Lu J., Notkins A.L.;
RT   "cDNA sequence and genomic organization of mouse secretin.";
RL   Biochem. Biophys. Res. Commun. 200:1066-1071(1994).
CC   -!- FUNCTION: Stimulates formation of NaHCO(3)-rich pancreatic juice
CC       and secretion of NaHCO(3)-rich bile and inhibits HCl production by
CC       the stomach.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the glucagon family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U07568; AAA18453.1; -.
DR   EMBL; X73580; CAA51982.1; -.
DR   PIR; JC2202; JC2202.
DR   MGD; MGI:99466; Sct.
```

```
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 1.
DR   SMART; SM00070; GLUCA; 1.
DR   PROSITE; PS00260; GLUCAGON; 1.
KW   Glucagon family; Hormone; Amidation;
KW   Cleavage on pair of basic residues; Signal.
FT   SIGNAL       1     22       BY SIMILARITY.
FT   PEPTIDE      32    58       SECRETIN (BY SIMILARITY).
FT   MOD_RES      58    58       AMIDATION (G-59 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   133 AA;  14914 MW;  9B69CBCF74CA9709 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 133;
  Best Local Similarity   80.0%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        129 TRPPR 133


RESULT 19
VEGH_ORFN2
ID   VEGH_ORFN2     STANDARD;      PRT;   133 AA.
AC   P52584;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Vascular endothelial growth factor homolog precursor.
GN   A2R.
OS   Orf virus (strain NZ2) (OV NZ-2).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Parapoxvirus.
OX   NCBI_TaxID=10259;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94076465; PubMed=8254780;
RA   Lyttle D.J., Fraser K.M., Fleming S.B., Mercer A.A., Robinson A.J.;
RT   "Homologs of vascular endothelial growth factor are encoded by the
RT   poxvirus orf virus.";
RL   J. Virol. 68:84-92(1994).
CC   -!- FUNCTION: INDUCES ENDOTHELIAL PROLIFERATION.
CC   -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC   -!- SIMILARITY: Belongs to the PDGF/VEGF growth factor family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S67520; AAB29220.2; -.
DR   HSSP; P15692; 1VPP.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR000072; PD_growth_factor.
DR   Pfam; PF00341; PDGF; 1.
DR   PRINTS; PR00438; GFCYSKNOT.
DR   ProDom; PD001629; PD_growth_factor; 1.
DR   SMART; SM00141; PDGF; 1.
DR   PROSITE; PS00249; PDGF_1; 1.
DR   PROSITE; PS50278; PDGF_2; 1.
KW   Mitogen; Growth factor; Glycoprotein; Signal.
FT   SIGNAL       1     20       POTENTIAL.
FT   CHAIN        21    133      VASCULAR ENDOTHELIAL GROWTH FACTOR
FT                               HOMOLOG.
FT   DISULFID     36     78      BY SIMILARITY.
FT   DISULFID     67    112      BY SIMILARITY.
FT   DISULFID     71    114      BY SIMILARITY.
FT   DISULFID     61     61      INTERCHAIN (BY SIMILARITY).
FT   DISULFID     70     70      INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD     85     85      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   133 AA;  14715 MW;  917C0F6883030C39 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 133;
  Best Local Similarity   80.0%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        126 TRPPR 130


RESULT 20
RT12_DROME
ID   RT12_DROME     STANDARD;      PRT;   140 AA.
AC   P10735; Q9V3R2;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   40S ribosomal protein S12, mitochondrial precursor (MT-RPS12)
DE   (Technical knockout locus protein).
GN   TKO OR EG:BACH59J11.1 OR CG7925.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88027001; PubMed=3117373;
RA   Royden C.S., Pirrotta V., Jan L.Y.;
RT   "The tko locus, site of a behavioral mutation in D. melanogaster,
RT   codes for a protein homologous to prokaryotic ribosomal protein
RT   S12.";
RL   Cell 51:165-173(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Oregon-R;
RX   MEDLINE=20196011; PubMed=10731137;
RA   Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
RA   Barrell B.G., Ferraz C., Vidal S., Brun C., Demailles J., Cadieu E.,
RA   Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,
RA   Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
RA   Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
RA   Modolell J., Peter A., Schoettler P., Werner M., Mourkioti F.,
RA   Beinert N., Dowe G., Schaefer U., Jaeckle H., Bucheton A.,
RA   Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,
RA   McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
RA   Glover D.M.;
RT   "From sequence to chromosome: the tip of the X chromosome of D.
RT   melanogaster.";
RL   Science 287:2220-2222(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
```

```
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- SUBCELLULAR LOCATION: Mitochondrial.
CC   -!- MISCELLANEOUS: MUTATION OF TKO CAUSES BEHAVIORAL MUTATION ("BANG
CC       SENSITIVITY" = TEMPORARILY PARALYSIS IN RESPONSE TO A PHYSICAL
CC       JOLT).
CC   -!- SIMILARITY: Belongs to the S12P family of ribosomal proteins.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M19494; AAA28935.1; -.
DR   EMBL; AL133505; CAB65841.1; -.
DR   EMBL; AE003424; AAF45781.1; -.
DR   PIR; A29622; A29622.
DR   FlyBase; FBgn0003714; tko.
DR   GO; GO:0008049; P:male courtship behavior; IMP.
DR   GO; GO:0007638; P:mechanosensory behavior; IMP.
DR   GO; GO:0009592; P:perception of sound; IMP.
DR   GO; GO:0009612; P:response to mechanical stimulus; IMP.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR006032; Ribosomal_S12_23.
DR   InterPro; IPR005679; Ribosomal_S12b/c.
DR   Pfam; PF00164; Ribosomal_S12; 1.
DR   PRINTS; PR01034; RIBOSOMALS12.
DR   ProDom; PD000576; Ribosomal_S12_23; 1.
DR   TIGRFAMs; TIGR00981; rpsL_bact; 1.
DR   PROSITE; PS00055; RIBOSOMAL_S12; 1.
KW   Ribosomal protein; Mitochondrion; Transit peptide.
FT   TRANSIT       1     30       MITOCHONDRION (POTENTIAL).
FT   CHAIN        31    140       40S RIBOSOMAL PROTEIN S12.
SQ   SEQUENCE   140 AA;  15532 MW;  8187A7312C529F31 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 140;
  Best Local Similarity   80.0%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db         45 TRPPR 49


RESULT 21
KGUA_SYNEL
ID   KGUA_SYNEL     STANDARD;      PRT;   191 AA.
AC   Q8DMQ7;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Guanylate kinase (EC 2.7.4.8) (GMP kinase).
GN   GMK OR TLL0054.
OS   Synechococcus elongatus (Thermosynechococcus elongatus).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX   NCBI_TaxID=32046;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BP-1;
RX   MEDLINE=22225144; PubMed=12240834;
RA   Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA   Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA   Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the thermophilic cyanobacterium
RT   Thermosynechococcus elongatus BP-1.";
RL   DNA Res. 9:123-130(2002).
CC   -!- FUNCTION: Essential for recycling GMP and indirectly, cGMP.
CC   -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the guanylate kinase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP005369; BAC07607.1; -.
DR   HAMAP; MF_00328; -; 1.
DR   InterPro; IPR008144; Guanylate_kin.
DR   InterPro; IPR008145; Guanylt/Ca.
DR   Pfam; PF00625; Guanylate_kin; 1.
DR   SMART; SM00072; GuKc; 1.
DR   PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR   PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
KW   Transferase; Kinase; ATP-binding; Complete proteome.
FT   NP_BIND      16     23       ATP (BY SIMILARITY).
SQ   SEQUENCE   191 AA;  20990 MW;  70E6C15768D73D51 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 191;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db         44 TRPPR 48


RESULT 22
VCO7_ADE04
ID   VCO7_ADE04     STANDARD;      PRT;   193 AA.
AC   Q96831;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Major core protein precursor (Protein VII) (pVII).
GN   PVII.
OS   Human adenovirus type 4.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=28280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Isolate RI-6;
RA   Tarassishin L., Szawlowski P.W.S., McLay J., Russell W.C.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ---------------------------------------------------------------------
DR   EMBL; U70921; AAC83411.1; -.
DR   InterPro; IPR004912; Adeno_VII.
DR   Pfam; PF03228; Adeno_VII; 1.
FT   PROPEP          1     24       BY SIMILARITY.
FT   CHAIN          25    193       MAJOR CORE PROTEIN.
FT   SITE           24     25       CLEAVAGE (BY ADENOVIRUS PROTEASE)
FT                                  (POTENTIAL).
SQ   SEQUENCE   193 AA;  21358 MW;  43137E07DB379DD0 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 193;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        188 TRPPR 192


RESULT 23
VCO7_ADE02
ID   VCO7_ADE02     STANDARD;      PRT;   198 AA.
AC   P03266; P12542;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Major core protein precursor (Protein VII) (pVII).
GN   PVII.
OS   Human adenovirus type 2, and
OS   Human adenovirus type 5.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=10515, 28285;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human adenovirus type 2;
RX   MEDLINE=85054835; PubMed=6094534;
RA   Alestroem P., Akusjaervi G., Lager M., Yeh-Kai L., Pettersson U.;
RT   "Genes encoding the core proteins of adenovirus type 2.";
RL   J. Biol. Chem. 259:13980-13985(1984).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human adenovirus type 2;
RX   MEDLINE=83221511; PubMed=6574459;
RA   Sung M.T., Cao T.M., Coleman R.T., Budelier K.A.;
RT   "Gene and protein sequences of adenovirus protein VII, a hybrid basic
RT   chromosomal protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:2902-2906(1983).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human adenovirus type 5;
RX   MEDLINE=92087470; PubMed=1727603;
RA   Chroboczek J., Bieber F., Jacrot B.;
RT   "The sequence of the genome of adenovirus type 5 and its comparison
RT   with the genome of adenovirus type 2.";
RL   Virology 186:280-285(1992).
RN   [4]
RP   SEQUENCE OF 1-136 FROM N.A.
RC   SPECIES=Human adenovirus type 5;
RX   MEDLINE=89137988; PubMed=3224820;
RA   Neumann R., Chroboczek J., Jacrot B.;
RT   "Determination of the nucleotide sequence for the penton-base gene of
RT   human adenovirus type 5.";
RL   Gene 69:153-157(1988).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J01917; AAA92212.1; -.
DR   EMBL; M73260; AAA96408.1; ALT_SEQ.
DR   EMBL; M22141; AAA42520.1; -.
DR   PIR; C03837; FOAD72.
DR   PIR; PT0067; FOADH5.
DR   InterPro; IPR004912; Adeno_VII.
DR   Pfam; PF03228; Adeno_VII; 1.
KW   Core protein; Late protein.
FT   PROPEP          1     24
FT   CHAIN          25    198       MAJOR CORE PROTEIN.
FT   SITE           24     25       CLEAVAGE (BY ADENOVIRUS PROTEASE).
FT   CONFLICT      112    112       MISSING (IN REF. 2).
SQ   SEQUENCE   198 AA;  21992 MW;  7D5A8D426F08E952 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 198;
  Best Local Similarity   80.0%;  Pred. No. 1.8e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        193 TRPPR 197


RESULT 24
KGUA_ANASP
ID   KGUA_ANASP     STANDARD;      PRT;   199 AA.
AC   Q8Z0I7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Guanylate kinase (EC 2.7.4.8) (GMP kinase).
GN   GMK OR ALR0106.
OS   Anabaena sp. (strain PCC 7120).
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX   NCBI_TaxID=103690;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21595285; PubMed=11759840;
RA   Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA   Yasuda M., Tabata S.;
RT   "Complete genomic sequence of the filamentous nitrogen-fixing
RT   cyanobacterium Anabaena sp. strain PCC 7120.";
RL   DNA Res. 8:205-213(2001).
CC   -!- FUNCTION: Essential for recycling GMP and indirectly, cGMP.
CC   -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the guanylate kinase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP003581; BAB77630.1; -.
DR   PIR; AB1820; AB1820.
DR   HAMAP; MF_00328; -; 1.
DR   InterPro; IPR008144; Guanylate_kin.
DR   InterPro; IPR008145; Guanylt/Ca.
DR   Pfam; PF00625; Guanylate_kin; 1.
DR   SMART; SM00072; GuKc; 1.
DR   PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR   PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
KW   Transferase; Kinase; ATP-binding; Complete proteome.
FT   NP_BIND       27     34       ATP (BY SIMILARITY).
SQ   SEQUENCE   199 AA;  22099 MW;  58634498D484850D CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 199;
  Best Local Similarity   80.0%;  Pred. No. 1.8e+02;
```

```
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db         55 TRPPR 59


RESULT 25
KGUA_NEIMA
ID   KGUA_NEIMA     STANDARD;      PRT;   205 AA.
AC   Q9JT96;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Guanylate kinase (EC 2.7.4.8) (GMP kinase).
GN   GMK OR NMA1919.
OS   Neisseria meningitidis (serogroup A).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=65699;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX   MEDLINE=20222556; PubMed=10761919;
RA   Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA   Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA   Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA   Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA   Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA   Whitehead S., Spratt B.G., Barrell B.G.;
RT   "Complete DNA sequence of a serogroup A strain of Neisseria
RT   meningitidis Z2491.";
RL   Nature 404:502-506(2000).
CC   -!- FUNCTION: Essential for recycling GMP and indirectly, cGMP.
CC   -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the guanylate kinase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL162757; CAB85140.1; -.
DR   PIR; G81819; G81819.
DR   HSSP; P15454; 1GKY.
DR   HAMAP; MF_00328; -; 1.
DR   InterPro; IPR008144; Guanylate_kin.
DR   InterPro; IPR008145; Guanylt/Ca.
DR   Pfam; PF00625; Guanylate_kin; 1.
DR   SMART; SM00072; GuKc; 1.
DR   PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR   PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
KW   Transferase; Kinase; ATP-binding; Complete proteome.
FT   NP_BIND      14     21       ATP (BY SIMILARITY).
SQ   SEQUENCE   205 AA;  22530 MW;  04FA07E5450C9007 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 205;
  Best Local Similarity   80.0%;  Pred. No. 1.8e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db         42 TRPPR 46


RESULT 26
KGUA_NEIMB
ID   KGUA_NEIMB     STANDARD;      PRT;   205 AA.
AC   Q9JYB5;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Guanylate kinase (EC 2.7.4.8) (GMP kinase).
GN   GMK OR NMB1661.
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
CC   -!- FUNCTION: Essential for recycling GMP and indirectly, cGMP.
CC   -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the guanylate kinase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE002517; AAF42010.1; -.
DR   PIR; G81055; G81055.
DR   HSSP; P15454; 1GKY.
DR   TIGR; NMB1661; -.
DR   HAMAP; MF_00328; -; 1.
DR   InterPro; IPR008144; Guanylate_kin.
DR   InterPro; IPR008145; Guanylt/Ca.
DR   Pfam; PF00625; Guanylate_kin; 1.
DR   SMART; SM00072; GuKc; 1.
DR   PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR   PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
KW   Transferase; Kinase; ATP-binding; Complete proteome.
FT   NP_BIND      14     21       ATP (BY SIMILARITY).
SQ   SEQUENCE   205 AA;  22500 MW;  54EE07E545189008 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 205;
  Best Local Similarity   80.0%;  Pred. No. 1.8e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db         42 TRPPR 46


RESULT 27
Y132_NPVAC
ID   Y132_NPVAC     STANDARD;      PRT;   219 AA.
AC   P24730;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Hypothetical 25.1 kDa protein in PP34-EXO intergenic region (ORF 4).
OS   Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC   Nucleopolyhedrovirus.
OX   NCBI_TaxID=46015;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87311863; PubMed=3041026;
RA   Oellig C., Happ B., Mueller T., Doerfler W.;
RT   "Overlapping sets of viral RNAs reflect the array of polypeptides in
RT   the EcoRI J and N fragments (map positions 81.2 to 85.0) of the
RT   Autographa californica nuclear polyhedrosis virus genome.";
RL   J. Virol. 61:3048-3057(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C6;
RX   MEDLINE=94303173; PubMed=8030224;
RA   Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT   "The complete DNA sequence of Autographa californica nuclear
RT   polyhedrosis virus.";
RL   Virology 202:586-605(1994).
CC   -!- SIMILARITY: TO CORRESPONDING ORF IN OPMNPV.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M17548; AAA66806.1; -.
DR   EMBL; L22858; AAA66762.1; -.
DR   PIR; E72866; E72866.
KW   Hypothetical protein; Late protein.
SQ   SEQUENCE   219 AA;  25136 MW;  281E1625BE8F6A5E CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 219;
  Best Local Similarity   80.0%;  Pred. No. 2e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db         23 TKPPK 27


RESULT 28
PE55_LUCCU
ID   PE55_LUCCU     STANDARD;      PRT;   220 AA.
AC   Q95UE8; Q8MUP5;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Peritrophin-55 precursor.
OS   Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC   Calliphoridae; Lucilia.
OX   NCBI_TaxID=7375;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 20-57 AND 60-94, TISSUE SPECIFICITY,
RP   AND DEVELOPMENTAL STAGE.
RC   TISSUE=Larva;
RX   MEDLINE=22423280; PubMed=12535682;
RA   Tellam R.L., Vuocolo T., Eisemann C.H., Briscoe S., Riding G.A.,
RA   Elvin C.M., Pearson R.D.;
RT   "Identification of an immuno-protective mucin-like protein,
RT   peritrophin-55, from the peritrophic matrix of Lucilia cuprina
RT   larvae.";
RL   Insect Biochem. Mol. Biol. 33:239-252(2003).
CC   -!- FUNCTION: May bind oligosaccharide structures.
CC   -!- TISSUE SPECIFICITY: Larval peritrophic membrane.
CC   -!- DEVELOPMENTAL STAGE: Expressed in all 3 larval instars but not
CC       adults or eggs.
CC   -!- PTM: Glycosylated.
CC   -!- SIMILARITY: Contains 1 chitin-binding type-2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY055470; AAL15463.1; -.
DR   EMBL; AF515826; AAM55223.1; -.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   PROSITE; PS50940; CHIT_BIND_II; 1.
KW   Glycoprotein; Signal.
FT   SIGNAL       1     19
FT   CHAIN       20    220       PERITROPHIN-55.
FT   DOMAIN      33     95       CHITIN-BINDING TYPE-2.
FT   CARBOHYD    29     29       N-LINKED (GLCNAC. . .) (PROBABLE).
FT   CONFLICT    67     67       D -> H (IN REF. 1; AAL15463).
FT   CONFLICT    69     69       P -> F (IN REF. 1; AA SEQUENCE).
FT   CONFLICT   142    142       I -> P (IN REF. 1; AAM55223).
FT   CONFLICT   164    164       V -> G (IN REF. 1; AAM55223).
FT   CONFLICT   190    190       S -> P (IN REF. 1; AAM55223).
SQ   SEQUENCE   220 AA;  23535 MW;  88C74ED57F2ED7C7 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 220;
  Best Local Similarity   80.0%;  Pred. No. 2e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db        155 TKPPK 159


RESULT 29
NOCT_RAT
ID   NOCT_RAT       STANDARD;      PRT;   253 AA.
AC   Q9ET55;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Nocturnin (CCR4 protein homolog) (Fragment).
GN   CCRN4L OR NOC OR CCR4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Retina;
RA   Wang Y., Osterbur D.L., Green C.B., Besharse J.C.;
RT   "Mammalian homologs of Xenopus nocturnin: conservation of structure
RT   and circadian regulation.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Component of the circadian clock or downstream effector
CC       of clock function. Exhibits a high amplitude circadian rhythm with
CC       maximal levels in early evening. In constant darkness or constant
CC       light, the amplitude of the rhythm decreases (By similarity).
CC   -!- SIMILARITY: Belongs to the CCR4/nocturin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF199495; AAG01390.1; -.
DR   InterPro; IPR005135; Exo_endo_phos.
DR   Pfam; PF03372; Exo_endo_phos; 1.
KW   Biological rhythms.
FT   NON_TER      1      1
FT   NON_TER    253    253
SQ   SEQUENCE   253 AA;  28662 MW;  C63566735BCC432E CRC64;
```

```
  Query Match             89.7%;  Score 26;  DB 1;  Length 253;
  Best Local Similarity   80.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db         44 TRPPR 48


RESULT 30
TRUA_XYLFA
ID   TRUA_XYLFA     STANDARD;      PRT;   257 AA.
AC   Q9PDK6;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   tRNA pseudouridine synthase A (EC 4.2.1.70) (Pseudouridylate synthase
DE   I) (Pseudouridine synthase I) (Uracil hydrolyase).
GN   TRUA OR XF1373.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9a5c;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
CC   -!- FUNCTION: Formation of pseudouridine at positions 38, 39 and 40 in
CC       the anticodon stem and loop of transfer RNAs (By similarity).
CC   -!- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate = pseudouridine
CC       5'-phosphate + H(2)O.
CC   -!- SIMILARITY: Belongs to the pseudouridine synthase truA family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE003968; AAF84182.1; -.
DR   PIR; A82691; A82691.
DR   HAMAP; MF_00171; -; 1.
DR   InterPro; IPR001406; PseudoU_synth_1.
DR   Pfam; PF01416; PseudoU_synth_1; 2.
DR   TIGRFAMs; TIGR00071; hisT_truA; 1.
KW   Lyase; tRNA processing; Complete proteome.
FT   ACT_SITE     53      53       BY SIMILARITY.
SQ   SEQUENCE   257 AA;  28638 MW;  C717E0D287C7D3F3 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 257;
  Best Local Similarity   80.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db         70 TRPPR 74


RESULT 31
TRUA_XYLFT
ID   TRUA_XYLFT     STANDARD;      PRT;   257 AA.
AC   Q87DS1;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   tRNA pseudouridine synthase A (EC 4.2.1.70) (Pseudouridylate synthase
DE   I) (Pseudouridine synthase I) (Uracil hydrolyase).
GN   TRUA OR PD0610.
OS   Xylella fastidiosa (strain Temecula1 / ATCC 700964).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xylella.
OX   NCBI_TaxID=183190;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22421331; PubMed=12533478;
RA   Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA   Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA   Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA   Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorry H., Tsai S.M.,
RA   Carrer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA   Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA   Marino C.L., Giglioti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA   Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA   da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA   Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sassaki F.T., Sena J.A.D.,
RA   de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA   Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA   Kitajima J.P.;
RT   "Comparative analyses of the complete genome sequences of Pierce's
RT   disease and citrus variegated chlorosis strains of Xylella
RT   fastidiosa.";
RL   J. Bacteriol. 185:1018-1026(2003).
CC   -!- FUNCTION: Formation of pseudouridine at positions 38, 39 and 40 in
CC       the anticodon stem and loop of transfer RNAs (By similarity).
CC   -!- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate = pseudouridine
CC       5'-phosphate + H(2)O.
CC   -!- SIMILARITY: Belongs to the pseudouridine synthase truA family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE012555; AAO28482.1; ALT_INIT.
DR   HAMAP; MF_00171; -; 1.
DR   InterPro; IPR001406; PseudoU_synth_1.
DR   Pfam; PF01416; PseudoU_synth_1; 2.
DR   TIGRFAMs; TIGR00071; hisT_truA; 1.
KW   Lyase; tRNA processing; Complete proteome.
FT   ACT_SITE     53      53       BY SIMILARITY.
SQ   SEQUENCE   257 AA;  28664 MW;  F6D5A76D9C2A88EC CRC64;
```

```
  Query Match             89.7%;  Score 26;  DB 1;  Length 257;
  Best Local Similarity   80.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db         70 TRPPR 74


RESULT 32
YQOB_CAEEL
ID   YQOB_CAEEL     STANDARD;      PRT;   304 AA.
AC   Q09300;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Hypothetical 33.0 kDa protein EEED8.11 in chromosome II precursor.
GN   EEED8.11.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Chissoe S.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: SOME, TO C.ELEGANS R13F6.2 AND R13F6.8.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U23484; AAC46771.1; -.
DR   PIR; T15922; T15922.
DR   WormPep; EEED8.11; CE01884.
DR   InterPro; IPR001304; Lectin_C.
DR   SMART; SM00034; CLECT; 1.
KW   Hypothetical protein; Signal.
FT   SIGNAL       1     19       POTENTIAL.
FT   CHAIN       20    304       HYPOTHETICAL PROTEIN EEED8.11.
FT   DOMAIN      64     92       POLY-THR.
SQ   SEQUENCE   304 AA;  32982 MW;  60C223B88F534151 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 304;
  Best Local Similarity   80.0%;  Pred. No. 2.7e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db         51 TKPPK 55


RESULT 33
YB04_AQUAE
ID   YB04_AQUAE     STANDARD;      PRT;   350 AA.
AC   O67189;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein AQ_1104.
GN   AQ_1104.
OS   Aquifex aeolicus.
OC   Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX   NCBI_TaxID=63363;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VF5;
RX   MEDLINE=98196666; PubMed=9537320;
RA   Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA   Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA   Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT   "The complete genome of the hyperthermophilic bacterium Aquifex
RT   aeolicus.";
RL   Nature 392:353-358(1998).
CC   -!- SIMILARITY: SOME, TO R.PROWAZEKII RP189.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000723; AAC07153.1; -.
DR   PIR; B70395; B70395.
DR   InterPro; IPR008921; Pol_clamp_load_C.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   350 AA;  40693 MW;  2949E786E1DAC2F9 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 350;
  Best Local Similarity   80.0%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db         14 TKPPK 18


RESULT 34
TRMU_RICCN
ID   TRMU_RICCN     STANDARD;      PRT;   370 AA.
AC   Q92IL0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
DE   (EC 2.1.1.61).
GN   TRMU OR RC0410.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
CC       homocysteine + tRNA containing 5-methylaminomethyl-2-
CC       thiouridylate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the trmU family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE008605; AAL02948.1; -.
DR   PIR; B97751; B97751.
DR   HAMAP; MF_00144; -; 1.
DR   InterPro; IPR004506; TrmU.
DR   Pfam; PF03054; tRNA_Me_trans; 1.
```

```
DR   TIGRFAMs; TIGR00420; trmU; 1.
KW   Transferase; Methyltransferase; tRNA processing; Complete proteome.
SQ   SEQUENCE   370 AA;  40915 MW;  1884AC815E730CE3 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 370;
  Best Local Similarity   80.0%;  Pred. No. 3.3e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        318 TRPPR 322


RESULT 35
RT31_HUMAN
ID   RT31_HUMAN     STANDARD;      PRT;   395 AA.
AC   Q92665; Q8WTV8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   28S ribosomal protein S31, mitochondrial precursor (S31mt) (MRP-S31)
DE   (Imogen 38).
GN   MRPS31 OR IMOGN38.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RA   Hutton J.C., Roep B.O.;
RT   "Human Imogen 38. T-cell and antibody responses in newly diagnosed
RT   diabetic subjects.";
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   IDENTIFICATION.
RX   MEDLINE=21276436; PubMed=11279123;
RA   Koc E.C., Burkhart W., Blackburn K., Moseley A., Spremulli L.L.;
RT   "The small subunit of the mammalian mitochondrial ribosome:
RT   identification of the full complement of ribosomal proteins present.";
RL   J. Biol. Chem. 276:19363-19374(2001).
CC   -!- SUBUNIT: Component of the mitochondrial ribosome small subunit
CC       (28S) which comprises a 12S rRNA and about 30 distinct proteins.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z68747; CAA92951.1; -.
DR   EMBL; BC022045; AAH22045.1; -.
DR   Genew; HGNC:16632; MRPS31.
DR   GO; GO:0005739; C:mitochondrion; TAS.
KW   Ribosomal protein; Mitochondrion; Transit peptide.
FT   TRANSIT       1     65       MITOCHONDRION (POTENTIAL).
FT   CHAIN        66    395       28S RIBOSOMAL PROTEIN S31.
FT   CONFLICT     80     80       I -> T (IN REF. 1).
FT   CONFLICT    132    132       A -> T (IN REF. 1).
FT   CONFLICT    279    279       D -> N (IN REF. 1).
SQ   SEQUENCE   395 AA;  45300 MW;  E9410F46C94C6F3D CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 395;
  Best Local Similarity   80.0%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db        118 TKPPK 122


RESULT 36
TRMU_RHILO
ID   TRMU_RHILO     STANDARD;      PRT;   396 AA.
AC   Q98HL0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
DE   (EC 2.1.1.61).
GN   TRMU OR MLR2824.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
CC       homocysteine + tRNA containing 5-methylaminomethyl-2-
CC       thiouridylate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the trmU family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AP003000; BAB49856.1; ALT_INIT.
DR   HAMAP; MF_00144; -; 1.
DR   InterPro; IPR004506; TrmU.
DR   Pfam; PF03054; tRNA_Me_trans; 1.
DR   TIGRFAMs; TIGR00420; trmU; 1.
KW   Transferase; Methyltransferase; tRNA processing; Complete proteome.
SQ   SEQUENCE   396 AA;  42537 MW;  66F118AA897E3086 CRC64;
```

```
  Query Match             89.7%;  Score 26;  DB 1;  Length 396;
  Best Local Similarity   80.0%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        320 TRPPR 324


RESULT 37
TRMU_BRUME
ID   TRMU_BRUME     STANDARD;      PRT;   398 AA.
AC   Q8YIL6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
DE   (EC 2.1.1.61).
GN   TRMU OR BMEI0428.
OS   Brucella melitensis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Brucellaceae; Brucella.
OX   NCBI_TaxID=29459;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=16M / ATCC 23456 / Biotype 1;
RX   MEDLINE=20020109; PubMed=11756688;
RA   DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T.,
RA   Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA   Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA   Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA   Haselkorn R., Kyrpides N., Overbeek R.;
RT   "The genome sequence of the facultative intracellular pathogen
RT   Brucella melitensis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
CC       homocysteine + tRNA containing 5-methylaminomethyl-2-
CC       thiouridylate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the trmU family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE009485; AAL51609.1; -.
DR   PIR; AF3305; AF3305.
DR   HAMAP; MF_00144; -; 1.
DR   InterPro; IPR004506; TrmU.
DR   Pfam; PF03054; tRNA_Me_trans; 1.
DR   TIGRFAMs; TIGR00420; trmU; 1.
KW   Transferase; Methyltransferase; tRNA processing; Complete proteome.
SQ   SEQUENCE   398 AA;  43175 MW;  93687D1D8BAB9ED3 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 398;
  Best Local Similarity   80.0%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        322 TRPPR 326


RESULT 38
TRMU_BRUSU
ID   TRMU_BRUSU     STANDARD;      PRT;   398 AA.
AC   Q8CY38;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
DE   (EC 2.1.1.61).
GN   TRMU OR BR1591.
OS   Brucella suis.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Brucellaceae; Brucella.
OX   NCBI_TaxID=29461;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1330 / Biovar 1;
RX   MEDLINE=22247741; PubMed=12271122;
RA   Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA   Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
RA   Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA   Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
RA   Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA   Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT   "The Brucella suis genome reveals fundamental similarities between
RT   animal and plant pathogens and symbionts.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
CC       homocysteine + tRNA containing 5-methylaminomethyl-2-
CC       thiouridylate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the trmU family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE014453; AAN30496.1; -.
DR   TIGR; BR1591; -.
DR   HAMAP; MF_00144; -; 1.
DR   InterPro; IPR004506; TrmU.
DR   Pfam; PF03054; tRNA_Me_trans; 1.
DR   TIGRFAMs; TIGR00420; trmU; 1.
KW   Transferase; Methyltransferase; tRNA processing; Complete proteome.
SQ   SEQUENCE   398 AA;  43228 MW;  937367AC3BAB9ED3 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 398;
  Best Local Similarity   80.0%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        322 TRPPR 326


RESULT 39
VGLD_HSVEA
ID   VGLD_HSVEA     STANDARD;      PRT;   402 AA.
AC   P24872;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycoprotein D precursor (Glycoprotein 17/18).
GN   GD OR GP17/18 OR 72.
OS   Equine herpesvirus type 1 (strain AB1) (EHV-1).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=10328;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92268882; PubMed=1316942;
RA   Elton D.M., Halliburton I.W., Killington R.A., Meredith D.M.,
RA   Bonass W.A.;
RT   "Identification of the equine herpesvirus type 1 glycoprotein 17/18
RT   as a homologue of herpes simplex virus glycoprotein D.";
```

```
RL   J. Gen. Virol. 73:1227-1233(1992).
RN   [2]
RP   SEQUENCE OF 242-402 FROM N.A.
RX   MEDLINE=91276272; PubMed=1647359;
RA   Elton D.M., Halliburton I.W., Killington R.A., Meredith D.M.,
RA   Bonass W.A.;
RT   "Sequence analysis of the 4.7-kb BamHI-EcoRI fragment of the equine
RT   herpesvirus type-1 short unique region.";
RL   Gene 101:203-208(1991).
CC   -!- SIMILARITY: Belongs to the herpesviruses glycoprotein D family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M60946; AAA46087.1; -.
DR   EMBL; M36299; AAA66546.1; -.
DR   InterPro; IPR002896; Herpes_glycop_D.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF01537; Herpes_glycop_D; 1.
KW   Glycoprotein; Transmembrane; Signal.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    402       GLYCOPROTEIN D.
FT   DOMAIN       31    355       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    356    372       POTENTIAL.
FT   DOMAIN      373    402       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     53     53       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    297    297       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    346    346       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   402 AA;  45211 MW;  78A0593232D0238C CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 402;
  Best Local Similarity   80.0%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db        337 TKPPK 341


RESULT 40
CCKR_HUMAN
ID   CCKR_HUMAN     STANDARD;      PRT;   428 AA.
AC   P32238;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Cholecystokinin type A receptor (CCK-A receptor) (CCK-AR).
GN   CCKAR OR CCKRA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Gall bladder;
RX   MEDLINE=93277552; PubMed=8503909;
RA   Ulrich C.D., Ferber I., Holicky E., Hadac E., Buell G.,
RA   Miller L.J.;
RT   "Molecular cloning and functional expression of the human gallbladder
RT   cholecystokinin A receptor.";
RL   Biochem. Biophys. Res. Commun. 193:204-211(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93343941; PubMed=8343165;
RA   Wank S.A., de Weerth A., Pisegna J.R., Huppi K.;
RT   "Molecular cloning, functional expression and chromosomal
RT   localization of the human cholecystokinin type A receptor.";
RL   Biochem. Biophys. Res. Commun. 194:811-818(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96029343; PubMed=7557108;
RA   Miller L.J., Holicky E.L., Ulrich C.D., Wieben E.D.;
RT   "Abnormal processing of the human cholecystokinin receptor gene in
RT   association with gallstones and obesity.";
RL   Gastroenterology 109:1375-1380(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Peripheral blood leukocytes;
RX   MEDLINE=20145045; PubMed=10682840;
RA   Funakoshi A., Miyasaka K., Matsumoto H., Yamamori S., Takiguchi S.,
RA   Kataoka K., Takata Y., Matsusue K., Kono A., Shimokata H.;
RT   "Gene structure of human cholecystokinin (CCK) type-A receptor: body
RT   fat content is related to CCK type-A receptor gene promoter
RT   polymorphism.";
RL   FEBS Lett. 466:264-266(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Stomach;
RA   Kopatz S.A., Aronstam R.S., Sharma S.V.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for cholecystokinin. Has a 1000 fold affinity
CC       for CCK rather than for gastrin. It modulates feeding and
CC       dopamine-induced behavior in the central and peripheral nervous
CC       system. This receptor mediates its action by association with G
CC       proteins that activate a phosphatidylinositol-calcium second
CC       messenger system.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L13605; AAA35659.1; -.
DR   EMBL; L19315; AAA02819.1; -.
DR   EMBL; U23430; AAA91123.1; -.
DR   EMBL; U23427; AAA91123.1; JOINED.
DR   EMBL; U23428; AAA91123.1; JOINED.
DR   EMBL; U23429; AAA91123.1; JOINED.
DR   EMBL; D85606; BAA90879.1; -.
DR   EMBL; AY322549; AAP84362.1; -.
DR   PIR; JN0692; JN0692.
DR   PDB; 1D6G; 17-NOV-99.
DR   PDB; 1HZN; 25-APR-01.
DR   Genew; HGNC:1570; CCKAR.
DR   MIM; 118444; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004951; F:cholecystokinin receptor activity; TAS.
DR   GO; GO:0007204; P:cytosolic calcium ion concentration elevation; TAS.
DR   GO; GO:0007586; P:digestion; TAS.
DR   GO; GO:0007631; P:feeding behavior; TAS.
DR   GO; GO:0007584; P:response to nutrients; TAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein;
KW   Palmitate; 3D-structure.
FT   DOMAIN        1     41       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     42     67       1 (POTENTIAL).
FT   DOMAIN       68     77       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     78    104       2 (POTENTIAL).
FT   DOMAIN      105    115       EXTRACELLULAR (POTENTIAL).
```

```
FT   TRANSMEM     116     137       3 (POTENTIAL).
FT   DOMAIN       138     157       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     158     178       4 (POTENTIAL).
FT   DOMAIN       179     210       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     211     234       5 (POTENTIAL).
FT   DOMAIN       235     313       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     314     334       6 (POTENTIAL).
FT   DOMAIN       335     349       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     350     373       7 (POTENTIAL).
FT   DOMAIN       374     428       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      10      10       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      24      24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     190     190       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     114     196       BY SIMILARITY.
FT   LIPID        387     387       S-palmitoyl cysteine (By similarity).
SQ   SEQUENCE   428 AA;  47841 MW;  A6E8FABDA805E610 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 428;
  Best Local Similarity   80.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        274 TRPPR 278


RESULT 41
NOCT_MOUSE
ID   NOCT_MOUSE     STANDARD;      PRT;   429 AA.
AC   O35710; Q9QZG9;
DT   15-JUL-1999 (Rel. 38, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Nocturnin (CCR4 protein homolog).
GN   CCRN4L OR NOC OR CCR4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Brain;
RX   MEDLINE=99453012; PubMed=10521507;
RA   Dupressoir A., Barbot W., Loireau M.P., Heidmann T.;
RT   "Characterization of a mammalian gene related to the yeast CCR4
RT   general transcription factor and revealed by transposon insertion.";
RL   J. Biol. Chem. 274:31068-31075(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Retina;
RX   MEDLINE=22944738; PubMed=11394964;
RA   Wang Y., Osterbur D.L., Megaw P.L., Tosini G., Fukuhara C.,
RA   Green C.B., Besharse J.C.;
RT   "Rhythmic expression of Nocturnin mRNA in multiple tissues of the
RT   mouse.";
RL   BMC Dev. Biol. 1:9-9(2001).
RN   [3]
RP   SEQUENCE OF 62-429 FROM N.A.
RC   STRAIN=DBA/2J; TISSUE=Liver;
RX   MEDLINE=97190339; PubMed=9038221;
RA   Puech A., Dupressoir A., Loireau M.P., Mattei M.-G., Heidmann T.;
RT   "Characterization of two age-induced intracisternal A-particle-related
RT   transcripts in the mouse liver. Transcriptional read-through into an
RT   open reading frame with similarities to the yeast ccr4 transcription
RT   factor.";
RL   J. Biol. Chem. 272:5995-6003(1997).
CC   -!- FUNCTION: Component of the circadian clock or downstream effector
CC       of clock function. Exhibits a high amplitude circadian rhythm with
CC       maximal levels in early evening. In constant darkness or constant
CC       light, the amplitude of the rhythm decreases (By similarity).
CC   -!- SIMILARITY: Belongs to the CCR4/nocturin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF183960; AAD56547.1; -.
DR   EMBL; AF199491; AAG01384.1; -.
DR   EMBL; U70139; AAB62717.1; ALT_FRAME.
DR   MGD; MGI:109382; Ccrn4l.
DR   InterPro; IPR005135; Exo_endo_phos.
DR   Pfam; PF03372; Exo_endo_phos; 1.
KW   Biological rhythms.
FT   CONFLICT     123     125       YQR -> LPA (IN REF. 3).
SQ   SEQUENCE   429 AA;  48300 MW;  CB9FB55D84E13942 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 429;
  Best Local Similarity   80.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        118 TRPPR 122


RESULT 42
NOCT_HUMAN
ID   NOCT_HUMAN     STANDARD;      PRT;   431 AA.
AC   Q9UK39; Q9HD93; Q9HD94; Q9HD95;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Nocturnin (CCR4 protein homolog).
GN   CCRN4L OR NOC OR CCR4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99453012; PubMed=10521507;
RA   Dupressoir A., Barbot W., Loireau M.P., Heidmann T.;
RT   "Characterization of a mammalian gene related to the yeast CCR4
RT   general transcription factor and revealed by transposon insertion.";
RL   J. Biol. Chem. 274:31068-31075(1999).
RN   [2]
RP   SEQUENCE OF 67-431 FROM N.A.
RA   Wang Y., Osterbur D.L., Green C.B., Besharse J.C.;
RT   "Mammalian homologs of Xenopus nocturnin: conservation of structure
RT   and circadian regulation.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Component of the circadian clock or downstream effector
CC       of clock function. Exhibits a high amplitude circadian rhythm with
CC       maximal levels in early evening. In constant darkness or constant
CC       light, the amplitude of the rhythm decreases (By similarity).
CC   -!- SIMILARITY: Belongs to the CCR4/nocturin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF183961; AAD56548.1; -.
DR   EMBL; AF199492; AAG01387.1; -.
DR   EMBL; AF199493; AAG01388.1; -.
DR   EMBL; AF199494; AAG01389.1; -.
DR   Genew; HGNC:14254; CCRN4L.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0003700; F:transcription factor activity; TAS.
```

```
DR   GO; GO:0006366; P:transcription from Pol II promoter; TAS.
DR   InterPro; IPR005135; Exo_endo_phos.
DR   Pfam; PF03372; Exo_endo_phos; 1.
KW   Biological rhythms.
FT   CONFLICT     69     69       T -> N (IN REF. 2).
FT   CONFLICT     77     77       G -> A (IN REF. 2).
FT   CONFLICT    266    266       A -> T (IN REF. 2; AAG01389).
FT   CONFLICT    341    341       N -> S (IN REF. 2).
SQ   SEQUENCE   431 AA;  48150 MW;  B61EF484E8D29AF5 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 431;
  Best Local Similarity   80.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        120 TRPPR 124


RESULT 43
ENGA_CLOPE
ID   ENGA_CLOPE     STANDARD;      PRT;   438 AA.
AC   Q8XJK1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   GTP-binding protein engA.
GN   ENGA OR CPE1755.
OS   Clostridium perfringens.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1502;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=13 / Type A;
RX   MEDLINE=21664373; PubMed=11792842;
RA   Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA   Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT   "Complete genome sequence of Clostridium perfringens, an anaerobic
RT   flesh-eater.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC   -!- FUNCTION: GTPase of unknown physiological role.
CC   -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
CC       proteins. EngA subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP003191; BAB81461.1; -.
DR   HAMAP; MF_00195; -; 1.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR005289; GTP-bindding_dom.
DR   InterPro; IPR006073; GTP1_OBG.
DR   InterPro; IPR002917; MMR_HSR1.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF01926; MMR_HSR1; 1.
DR   PRINTS; PR00326; GTP1OBG.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00382; AAA; 2.
DR   TIGRFAMs; TIGR00650; MG442; 2.
DR   TIGRFAMs; TIGR00231; small_GTP; 2.
KW   GTP-binding; Repeat; Complete proteome.
FT   NP_BIND      10     17       GTP 1 (POTENTIAL).
FT   NP_BIND      57     61       GTP 1 (POTENTIAL).
FT   NP_BIND     120    123       GTP 1 (POTENTIAL).
FT   NP_BIND     183    190       GTP 2 (POTENTIAL).
FT   NP_BIND     230    234       GTP 2 (POTENTIAL).
FT   NP_BIND     295    298       GTP 2 (POTENTIAL).
SQ   SEQUENCE   438 AA;  49586 MW;  9658B0BC98398396 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 438;
  Best Local Similarity   80.0%;  Pred. No. 4e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db        391 TKPPK 395


RESULT 44
SRMB_HAEIN
ID   SRMB_HAEIN     STANDARD;      PRT;   439 AA.
AC   P44701;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ATP-dependent RNA helicase srmB homolog.
GN   SRMB OR HI0422.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rd / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT   Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: RNA-DEPENDENT ATPASE ACTIVITY. PROBABLY INTERACTS
CC       WITH 23S RIBOSOMAL RNA (BY SIMILARITY).
CC   -!- SIMILARITY: Belongs to the DEAD box helicase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32725; AAC22078.1; -.
DR   PIR; H64066; H64066.
DR   HSSP; Q58083; 1HV8.
DR   TIGR; HI0422; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding; Complete proteome.
FT   NP_BIND      48     55       ATP (POTENTIAL).
FT   SITE        157    160       DEAD BOX.
SQ   SEQUENCE   439 AA;  49805 MW;  7543942CE35B20C4 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 439;
  Best Local Similarity   80.0%;  Pred. No. 4e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 TKPPR 5
              ||||:
Db        383 TKPPK 387


RESULT 45
VGLD_HSVEK
ID   VGLD_HSVEK     STANDARD;      PRT;   442 AA.
AC   P22484;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycoprotein D precursor (Glycoprotein 17/18).
GN   GD OR GP17/18 OR 72.
OS   Equine herpesvirus type 1 (strain Kentucky A) (EHV-1).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=10329;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082407; PubMed=1845821;
RA   Flowers C.C., Eastman E.M., O'Callaghan D.J.;
RT   "Sequence analysis of a glycoprotein D gene homolog within the unique
RT   short segment of the EHV-1 genome.";
RL   Virology 180:175-184(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92263758; PubMed=1316673;
RA   Colle C.F. III, Flowers C.C., O'Callaghan D.J.;
RT   "Open reading frames encoding a protein kinase, homolog of
RT   glycoprotein gX of pseudorabies virus, and a novel glycoprotein map
RT   within the unique short segment of equine herpesvirus type 1.";
RL   Virology 188:545-557(1992).
CC   -!- SIMILARITY: Belongs to the herpesviruses glycoprotein D family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M62923; AAA46081.1; -.
DR   EMBL; M86931; -; NOT_ANNOTATED_CDS.
DR   EMBL; M87497; AAA46073.1; ALT_INIT.
DR   PIR; A38518; VGBEEA.
DR   InterPro; IPR002896; Herpes_glycop_D.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF01537; Herpes_glycop_D; 1.
KW   Glycoprotein; Signal; Transmembrane.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    442       GLYCOPROTEIN D.
FT   DOMAIN       20    405       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    406    422       POTENTIAL.
FT   DOMAIN      423    442       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    103    103       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    111    111       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    347    347       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    396    396       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   442 AA;  49908 MW;  323CDCA9C9762F05 CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 442;
  Best Local Similarity   80.0%;  Pred. No. 4e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db        387 TKPPK 391


RESULT 46
VGLD_HSVEB
ID   VGLD_HSVEB     STANDARD;      PRT;   452 AA.
AC   P24379;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycoprotein D precursor (Glycoprotein 17/18).
GN   GD OR GP17/18 OR 72.
OS   Equine herpesvirus type 1 (strain Ab4p) (EHV-1), and
OS   Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=31520, 10330;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB4P;
RX   MEDLINE=92295566; PubMed=1318606;
RA   Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT   "The DNA sequence of equine herpesvirus-1.";
RL   Virology 189:304-316(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Kentucky D;
RX   MEDLINE=91108393; PubMed=2177089;
RA   Audonnet J.-C., Winslow J., Allen G., Paoletti E.;
RT   "Equine herpesvirus type 1 unique short fragment encodes
RT   glycoproteins with homology to herpes simplex virus type 1 gD, gI and
RT   gE.";
RL   J. Gen. Virol. 71:2969-2978(1990).
CC   -!- SIMILARITY: Belongs to the herpesviruses glycoprotein D family.
CC   -!- CAUTION: It is uncertain whether Met-1 or Met-51 is the initiator.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M86664; AAB02507.1; -.
DR   PIR; I36802; VGBEG3.
DR   InterPro; IPR002896; Herpes_glycop_D.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF01537; Herpes_glycop_D; 1.
KW   Glycoprotein; Signal; Transmembrane.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    452       GLYCOPROTEIN D.
FT   DOMAIN       20    405       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    406    422       POTENTIAL.
FT   DOMAIN      423    452       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    103    103       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    111    111       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    347    347       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    396    396       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   452 AA;  51099 MW;  CF51E914F7F2E9DC CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 452;
  Best Local Similarity   80.0%;  Pred. No. 4.1e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db        387 TKPPK 391


RESULT 47
ENGA_BIFLO
ID   ENGA_BIFLO     STANDARD;      PRT;   463 AA.
AC   Q8G6A8;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
```

```
DE   GTP-binding protein engA.
GN   ENGA OR BL0738.
OS   Bifidobacterium longum.
OC   Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC   Bifidobacteriaceae; Bifidobacterium.
OX   NCBI_TaxID=216816;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCC 2705;
RX   MEDLINE=22294977; PubMed=12381787;
RA   Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA   Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
RA   Pridmore R.D., Arigoni F.;
RT   "The genome sequence of Bifidobacterium longum reflects its adaptation
RT   to the human gastrointestinal tract.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
CC   -!- FUNCTION: GTPase of unknown physiological role.
CC   -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
CC       proteins. EngA subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE014696; AAN24555.1; -.
DR   HAMAP; MF_00195; -; 1.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR005289; GTP-bindding_dom.
DR   InterPro; IPR006073; GTP1_OBG.
DR   InterPro; IPR002917; MMR_HSR1.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF01926; MMR_HSR1; 1.
DR   PRINTS; PR00326; GTP1OBG.
DR   SMART; SM00382; AAA; 2.
DR   TIGRFAMs; TIGR00650; MG442; 2.
DR   TIGRFAMs; TIGR00231; small_GTP; 2.
KW   GTP-binding; Repeat; Complete proteome.
FT   NP_BIND      33     40       GTP 1 (POTENTIAL).
FT   NP_BIND      80     84       GTP 1 (POTENTIAL).
FT   NP_BIND     142    145       GTP 1 (POTENTIAL).
FT   NP_BIND     208    215       GTP 2 (POTENTIAL).
FT   NP_BIND     255    259       GTP 2 (POTENTIAL).
FT   NP_BIND     320    323       GTP 2 (POTENTIAL).
SQ   SEQUENCE   463 AA;  51186 MW;  7F986F8E041367AE CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 463;
  Best Local Similarity   80.0%;  Pred. No. 4.2e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        414 TRPPR 418


RESULT 48
EM55_HUMAN
ID   EM55_HUMAN     STANDARD;      PRT;   466 AA.
AC   Q00013;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   55 kDa erythrocyte membrane protein (p55) (Membrane protein,
DE   palmitoylated 1).
GN   MPP1 OR EMP55.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Reticulocytes;
RX   MEDLINE=91319732; PubMed=1713685;
RA   Ruff P., Speicher D.W., Husain-Chishti A.;
RT   "Molecular identification of a major palmitoylated erythrocyte
RT   membrane protein containing the src homology 3 motif.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:6595-6599(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93244792; PubMed=1301163;
RA   Metzenberg A.B., Gitschier J.;
RT   "The gene encoding the palmitoylated erythrocyte membrane protein,
RT   p55, originates at the CpG island 3' to the factor VIII gene.";
RL   Hum. Mol. Genet. 1:97-101(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBUNIT: Interacts with DLG5.
CC   -!- SUBCELLULAR LOCATION: Membrane-associated.
CC   -!- PTM: Extensively palmitoylated.
CC   -!- SIMILARITY: Belongs to the MAGUK family.
CC   -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC   -!- SIMILARITY: Contains 1 SH3 domain.
CC   -!- SIMILARITY: Contains 1 guanylate kinase-like domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M64925; AAA60059.1; -.
DR   EMBL; M87059; AAA60060.1; -.
DR   EMBL; U39611; AAD14835.1; -.
DR   EMBL; BC002392; AAH02392.1; -.
DR   HSSP; O14936; 1KWA.
DR   Genew; HGNC:7219; MPP1.
DR   MIM; 305360; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005624; C:membrane fraction; TAS.
DR   GO; GO:0004385; F:guanylate kinase activity; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR008144; Guanylate_kin.
DR   InterPro; IPR008145; Guanylt/Ca.
DR   InterPro; IPR001478; PDZ.
DR   InterPro; IPR001452; SH3.
DR   Pfam; PF00625; Guanylate_kin; 1.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF00018; SH3; 1.
DR   ProDom; PD000066; SH3; 1.
```

```
DR   SMART; SM00072; GuKc; 1.
DR   SMART; SM00228; PDZ; 1.
DR   SMART; SM00326; SH3; 1.
DR   PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR   PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
DR   PROSITE; PS50106; PDZ; 1.
DR   PROSITE; PS50002; SH3; 1.
KW   Membrane; Erythrocyte; SH3 domain; Lipoprotein; Palmitate;
KW   Polymorphism.
FT   DOMAIN       71    152       PDZ.
FT   DOMAIN      158    228       SH3.
FT   DOMAIN      280    466       GUANYLATE KINASE.
FT   VARIANT     448    448       E -> Q (in dbSNP:14092).
FT                                /FTId=VAR_011914.
SQ   SEQUENCE   466 AA;  52296 MW;  DC68AA68EF48A26E CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 466;
  Best Local Similarity   80.0%;  Pred. No. 4.2e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        318 TRPPR 322


RESULT 49
ENP6_HUMAN
ID   ENP6_HUMAN     STANDARD;      PRT;   484 AA.
AC   O75354; Q9UJD1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 6 (EC 3.6.1.6)
DE   (NTPDase6) (CD39 antigen-like 2).
GN   ENTPD6 OR CD39L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Keratinocytes;
RX   MEDLINE=98341119; PubMed=9676430;
RA   Chadwick B.P., Frischauf A.-M.;
RT   "The CD39-like gene family: identification of three new human members
RT   (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of
RT   the gene family from Drosophila melanogaster.";
RL   Genomics 50:357-367(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
CC   -!- FUNCTION: Might support glycosylation reactions in the Golgi
CC       apparatus and, when released from cells, might catalyze the
CC       hydrolysis of extracellular nucleotides. Hydrolyzes preferentially
CC       nucleoside 5'-diphosphates, nucleoside 5'-triphosphates are
CC       hydrolyzed only to a minor extent, there is no hydrolysis of
CC       nucleoside 5'-monophosphates. The order of activity with different
CC       substrates is GDP > IDP >> UDP = CDP >> ADP (By similarity).
CC   -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC       nucleotide + phosphate.
CC   -!- COFACTOR: Requires calcium and magnesium (By similarity).
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI. BUT ALSO
CC       OCCURS IN A SOLUBLE EXTRACELLULAR FORM (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: Expressed in most tissues.
CC   -!- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF039916; AAC39883.1; -.
DR   EMBL; AL035252; CAB41571.1; -.
DR   Genew; HGNC:3368; ENTPD6.
DR   MIM; 603160; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE_NEG.
KW   Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW   Signal-anchor; Golgi stack.
FT   DOMAIN        1     39       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     40     60       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       61    484       LUMENAL (POTENTIAL).
FT   CARBOHYD    220    220       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    284    284       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    138    138       V -> L (IN REF. 2).
FT   CONFLICT    202    202       E -> K (IN REF. 2).
SQ   SEQUENCE   484 AA;  53233 MW;  27334B290DB8D64C CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 484;
  Best Local Similarity   80.0%;  Pred. No. 4.4e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        121 TRPPR 125


RESULT 50
AMPA_AERPE
ID   AMPA_AERPE     STANDARD;      PRT;   524 AA.
AC   Q9Y935;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
DE   (LAP) (Leucyl aminopeptidase).
GN   PEPA OR APE2450.
OS   Aeropyrum pernix.
OC   Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC   Desulfurococcaceae; Aeropyrum.
OX   NCBI_TaxID=56636;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K1;
```

```
RX   MEDLINE=99310339; PubMed=10382966;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA   Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA   Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA   Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA   Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA   Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT   "Complete genome sequence of an aerobic hyper-thermophilic
RT   crenarchaeon, Aeropyrum pernix K1.";
RL   DNA Res. 6:83-101(1999).
CC   -!- FUNCTION: Presumably involved in the processing and regular
CC       turnover of intracellular proteins. Catalyzes the removal of
CC       unsubstituted N-terminal amino acids from various peptides (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-
CC       Xbb-, in which Xaa is preferably Leu, but may be other amino acids
CC       including Pro although not Arg or Lys, and Xbb may be Pro.
CC   -!- COFACTOR: Binds 2 manganese ions per subunit (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to peptidase family M17.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP000064; BAA81465.1; -.
DR   PIR; A72476; A72476.
DR   HSSP; P00727; 1LAP.
DR   MEROPS; M17.UPW; -.
DR   HAMAP; MF_00181; -; 1.
DR   InterPro; IPR000819; Peptidase_M17_C.
DR   InterPro; IPR008283; Peptidase_M17_N.
DR   Pfam; PF00883; Peptidase_M17; 1.
DR   Pfam; PF02789; Peptidase_M17_N; 1.
DR   PRINTS; PR00481; LAMNOPPTDASE.
DR   PROSITE; PS00631; CYTOSOL_AP; 1.
KW   Hydrolase; Aminopeptidase; Manganese; Complete proteome.
FT   METAL       280    280       MANGANESE 2 (BY SIMILARITY).
FT   METAL       285    285       MANGANESE 1 AND 2 (BY SIMILARITY).
FT   METAL       303    303       MANGANESE 2 (BY SIMILARITY).
FT   METAL       362    362       MANGANESE 1 (BY SIMILARITY).
FT   METAL       364    364       MANGANESE 1 AND 2 (BY SIMILARITY).
FT   ACT_SITE    292    292       POTENTIAL.
FT   ACT_SITE    366    366       POTENTIAL.
SQ   SEQUENCE   524 AA;  55609 MW;  0D6161A4D814C75D CRC64;

  Query Match             89.7%;  Score 26;  DB 1;  Length 524;
  Best Local Similarity   80.0%;  Pred. No. 4.8e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db         37 TRPPR 41


Search completed: March  3, 2004, 12:17:40
Job time : 14 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 3, 2004, 12:15:23 ; Search time 20 Seconds
(without alignments)
24.048 Million cell updates/sec

Title: US-09-871-974-2
Perfect score: 29
Sequence: 1 TKPPR 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	142	2	C82829	conserved hypothet
2	29	100.0	147	2	AH3157	conserved hypothet
3	29	100.0	191	2	T47369	hypothetical prote
4	29	100.0	201	2	S76026	hypothetical prote
5	29	100.0	222	2	G98129	hypothetical prote
6	29	100.0	226	2	T25360	hypothetical prote
7	29	100.0	278	2	S75883	hypothetical prote
8	29	100.0	283	2	T18299	hypothetical prote
9	29	100.0	350	2	AE3171	ATP-dependent DNA
10	29	100.0	351	2	T29369	hypothetical prote
11	29	100.0	370	2	AC0173	probable iron-sulf
12	29	100.0	375	2	D84715	probable protein k
13	29	100.0	410	2	S38238	hypothetical prote
14	29	100.0	415	2	F96499	hypothetical prote
15	29	100.0	415	2	C84829	hypothetical prote
16	29	100.0	442	2	S61165	hypothetical prote
17	29	100.0	493	2	S73752	hypothetical prote
18	29	100.0	575	2	T21775	hypothetical prote
19	29	100.0	604	2	S56027	hypothetical prote
20	29	100.0	615	2	JC7576	transcription fact
21	29	100.0	619	2	A43361	Ets-related transc
22	29	100.0	722	2	T22359	hypothetical prote
23	29	100.0	816	2	T45684	hypothetical prote
24	29	100.0	919	2	A41275	DNA ligase (ATP) (
25	29	100.0	941	2	T51135	ligand-gated chann
26	29	100.0	962	2	D86186	hypothetical prote
27	29	100.0	976	2	T51137	ionotropic glutama
28	29	100.0	1235	1	PWBYH	potassium transpor
29	29	100.0	1241	2	JU0466	potassium transpor
30	26	89.7	47	2	G81008	hypothetical prote
31	26	89.7	50	2	AH2135	hypothetical prote
32	26	89.7	66	2	C81143	hypothetical prote
33	26	89.7	68	2	T36774	hypothetical prote
34	26	89.7	75	2	AE1652	hypothetical prote
35	26	89.7	87	2	AD0288	conserved hypothet
36	26	89.7	88	2	T17560	hypothetical prote
37	26	89.7	100	2	F72587	hypothetical prote
38	26	89.7	133	2	JC2202	secretin precursor
39	26	89.7	133	2	B49530	vascular endotheli
40	26	89.7	140	2	A29622	ribosomal protein
41	26	89.7	140	2	T27059	hypothetical prote
42	26	89.7	144	2	T49706	hypothetical prote
43	26	89.7	148	2	A87466	hypothetical prote
44	26	89.7	153	2	S75561	hypothetical prote
45	26	89.7	154	2	A69267	transposase homolo
46	26	89.7	163	2	T49798	conserved hypothet
47	26	89.7	179	2	AB1974	hypothetical prote
48	26	89.7	183	2	H83561	hypothetical prote
49	26	89.7	195	2	S45627	acidic endoprotein
50	26	89.7	198	1	FOAD72	major core protein
51	26	89.7	198	2	T49814	hypothetical prote
52	26	89.7	199	2	AB1820	guanylate kinase [
53	26	89.7	205	2	G81819	guanylate kinase (
54	26	89.7	205	2	G81055	guanylate kinase N
55	26	89.7	219	2	E72866	AcOrf-132 protein
56	26	89.7	220	2	T41869	AcMNPV orf132 - Bo
57	26	89.7	230	2	T34541	hypothetical prote
58	26	89.7	232	2	G72508	hypothetical prote
59	26	89.7	257	2	A82691	tRNA pseudouridine
60	26	89.7	288	2	T23159	hypothetical prote
61	26	89.7	301	2	D90224	spermidine synthas
62	26	89.7	304	2	T15922	hypothetical prote
63	26	89.7	305	2	D83488	probable dihydrodi
64	26	89.7	320	2	E75255	hypothetical prote
65	26	89.7	330	2	H90100	putative transcrip
66	26	89.7	338	2	E75497	LuxA-related prote
67	26	89.7	338	2	PC4365	calcium-binding pr
68	26	89.7	350	2	S35631	genome polyprotein
69	26	89.7	350	2	B70395	hypothetical prote
70	26	89.7	353	2	AH3238	ATP-dependent DNA
71	26	89.7	359	1	H64688	site-specific DNA
72	26	89.7	359	2	F71827	type II DNA modifi
73	26	89.7	370	2	B97751	hypothetical prote
74	26	89.7	372	2	A53050	brefeldin A estera
75	26	89.7	398	2	AF3305	tRNA (5-methylamin
76	26	89.7	419	2	T39248	probable GTP bindi
77	26	89.7	428	2	JN0692	cholecystokinin ty
78	26	89.7	436	2	T36320	hypothetical prote
79	26	89.7	439	2	H64066	ATP-dependent RNA
80	26	89.7	442	1	VGBEEA	glycoprotein D pre
81	26	89.7	452	1	VGBEG3	glycoprotein D pre
82	26	89.7	466	2	A39599	55K erythrocyte me
83	26	89.7	496	2	S26402	homeotic protein H
84	26	89.7	524	2	A72476	probable cytosol a
85	26	89.7	526	2	T05681	phosphoinositide-s
86	26	89.7	530	2	A89783	oligopeptide trans
87	26	89.7	536	1	FOMVME	gag polyprotein -
88	26	89.7	546	2	C97554	hypothetical prote
89	26	89.7	549	2	T11688	phosphoinositide-s
90	26	89.7	551	2	T06777	phosphoinositide-s
91	26	89.7	553	2	T06771	phosphoinositide-s
92	26	89.7	565	2	T07424	phosphoinositide-s
93	26	89.7	579	2	AE2774	conserved hypothet
94	26	89.7	581	1	S71170	phosphoinositide-s
95	26	89.7	581	2	T50841	phosphoinositide-s
96	26	89.7	584	2	T49206	phosphoinositide-s
97	26	89.7	585	2	T07425	phosphoinositide-s
98	26	89.7	586	2	T50843	phospholipase C1 [
99	26	89.7	594	2	T06467	phosphoinositide-s
100	26	89.7	596	2	T07421	phosphoinositide-s
101	26	89.7	600	2	T06775	phosphoinositide-s
102	26	89.7	600	2	T06420	phosphoinositide-s

103	26	89.7	605	2	T50842	phospholipase C2 [
104	26	89.7	609	2	A43906	nuclear phosphopro
105	26	89.7	619	2	A36682	72K mitochondrial
106	26	89.7	624	2	T21072	hypothetical prote
107	26	89.7	666	2	B70803	hypothetical prote
108	26	89.7	744	2	A43353	ascites sialoglyco
109	26	89.7	747	2	S69557	hypothetical prote
110	26	89.7	759	2	T35749	probable formate d
111	26	89.7	784	2	AG2736	exoribonuclease [i
112	26	89.7	784	2	E97517	exoribonuclease RN
113	26	89.7	797	2	T33098	hypothetical prote
114	26	89.7	846	1	QQBEC3	HQRF1 protein - hu
115	26	89.7	905	2	C70758	probable cation tr
116	26	89.7	919	2	T29581	hypothetical prote
117	26	89.7	925	2	AE3628	nitrate reductase
118	26	89.7	961	1	TSHUP4	thrombospondin 4 p
119	26	89.7	965	2	T21073	hypothetical prote
120	26	89.7	975	2	T08606	protein phosphatas
121	26	89.7	983	2	T39902	translation Elonga
122	26	89.7	991	2	S43891	dna exoribonucleas
123	26	89.7	1057	2	S09112	hypothetical prote
124	26	89.7	1122	2	T42400	Eph receptor tyros
125	26	89.7	1127	2	T28435	proline dehydrogen
126	26	89.7	1158	2	E86327	protein F18O14.19
127	26	89.7	1160	2	T00272	hypothetical prote
128	26	89.7	1446	1	A45344	immediate-early pr
129	26	89.7	1460	1	EDBEIF	immediate-early pr
130	26	89.7	1630	2	A53577	ascites sialoglyco
131	26	89.7	2139	2	A44467	voltage-dependent
132	26	89.7	2143	2	JH0427	voltage-dependent
133	26	89.7	2166	2	S11339	calcium channel pr
134	26	89.7	2171	2	S05054	calcium channel al
135	26	89.7	2207	2	T42759	Munc13-3 protein -
136	26	89.7	2220	2	A45290	calcium channel pr
137	26	89.7	3033	1	GNWVJ8	genome polyprotein
138	26	89.7	4957	2	T03455	ALR protein - huma
139	26	89.7	5262	2	T03454	ALR protein - huma
140	26	89.7	5327	2	T13564	microtubule-associ
141	25	86.2	14	2	S65392	cytochrome-c oxida
142	25	86.2	69	1	B35537	cytochrome-c oxida
143	25	86.2	105	2	E72599	probable formylmet
144	25	86.2	111	1	L2HUWN	Ig lambda chain V-
145	25	86.2	126	2	H72614	hypothetical prote
146	25	86.2	133	2	E70690	hypothetical prote
147	25	86.2	168	2	C83505	probable outer mem
148	25	86.2	173	1	RUPSEO	rubredoxin II - Ps
149	25	86.2	181	2	S22990	zein, 27K - maize
150	25	86.2	199	2	S23635	gamma-coixin, 22K,

ALIGNMENTS

RESULT 1
C82829
conserved hypothetical protein XF0240 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: C82829
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequer
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82829
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-142 <SIM>
A;Cross-references: GB:AE003878; GB:AE003849; NID:g9105052; PIDN:AAF83053.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frol
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laic
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawas
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0240

Query Match 100.0%; Score 29; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 87 TKPPR 91

RESULT 2
AH3157
conserved hypothetical protein Atu4884 [imported] - Agrobacterium tumefaciens (strain
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AH3157
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McCle
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AH3157
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-147 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL45678.1; PID:g17743404; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu4884
A;Map position: linear chromosome

Query Match 100.0%; Score 29; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 2 TKPPR 6

RESULT 3
T47369
hypothetical protein F7M19.120 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: T47369
R;Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M
Mayer, K.F.X.
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24458
A;Accession: T47369
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-191 <NYA>
A;Cross-references: EMBL:AL138643
A;Experimental source: cultivar Columbia; BAC clone F7M19
C;Genetics:

A;Map position: 3
A;Note: F7M19.120

Query Match 100.0%; Score 29; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||||
Db 78 TKPPR 82

RESULT 4
S76026
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S76026
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76026
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-201 <KAN>
A;Cross-references: EMBL:D64006; GB:AB001339; NID:g1001291; PIDN:BAA10873.1; PID:g100138
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: conserved hypothetical protein YDR196c

Query Match 100.0%; Score 29; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||||
Db 3 TKPPR 7

RESULT 5
G98129
hypothetical protein AGR_L_19 [imported] - Agrobacterium tumefaciens (strain C58, Cereon
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C;Accession: G98129
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: G98129
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-222 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK88561.1; PID:g15158270; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_19
A;Map position: linear chromosome

Query Match 100.0%; Score 29; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||||
Db 77 TKPPR 81

RESULT 6
T25360
hypothetical protein T27E7.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25360
R;Cummings, P.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z20023
A;Accession: T25360
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-226 <WIL>
A;Cross-references: EMBL:Z82284; PIDN:CAB05287.1; GSPDB:GN00022; CESP:T27E7.1
A;Experimental source: clone T27E7
C;Genetics:
A;Gene: CESP:T27E7.1
A;Map position: 4
A;Introns: 20/1; 43/1; 101/3; 169/3

Query Match 100.0%; Score 29; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||||
Db 75 TKPPR 79

RESULT 7
S75883
hypothetical protein slr1169 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C;Accession: S75883
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocyst s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75883
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-278 <KAN>
A;Cross-references: EMBL:D90913; GB:AB001339; NID:g1653348; PIDN:BAA18342.1; PID:d1019
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 100.0%; Score 29; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||||
Db 240 TKPPR 244

RESULT 8
T18299
hypothetical protein - Entamoeba histolytica (fragment)
C;Species: Entamoeba histolytica
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C;Accession: T18299
R;Evans, J.D.
Proc. Natl. Acad. Sci. U.S.A. 92, 6518-6521, 1995
A;Title: Relatedness threshold for the production of female sexuals in colonies of a p
A;Reference number: Z18867; MUID:95327678; PMID:7604025
A;Accession: T18299
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-283 <EVA>
A;Cross-references: EMBL:L39933; NID:g6478875; PIDN:AAC41578.1; PID:g675517

Query Match 100.0%; Score 29; DB 2; Length 283;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 208 TKPPR 212

RESULT 9
AE3171
ATP-dependent DNA ligase Atu5097 [imported] - Agrobacterium tumefaciens (strain C58, Dup
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AE3171
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AE3171
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-350 <KUR>
A;Cross-references: GB:AE008687; PIDN:AAL45787.1; PID:g17743523; GSPDB:GN00188
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu5097
A;Genome: plasmid

Query Match 100.0%; Score 29; DB 2; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 2 TKPPR 6

RESULT 10
T29369
hypothetical protein ZC404.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29369
R;Bentley, D.; Le, T.T.
submitted to the EMBL Data Library, April 1996
A;Description: The sequence of C. elegans cosmid ZC404.
A;Reference number: Z20614
A;Accession: T29369
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-351 <BEN>
A;Cross-references: EMBL:U55363; PIDN:AAA97963.1; GSPDB:GN00023; CESP:ZC404.8
A;Experimental source: strain Bristol N2; clone ZC404
C;Genetics:
A;Gene: CESP:ZC404.8
A;Map position: 5
A;Introns: 17/2; 52/2; 73/2; 312/1

Query Match 100.0%; Score 29; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 199 TKPPR 203

RESULT 11
AC0173
probable iron-sulfur binding protein YPO1417 [imported] - Yersinia pestis (strain CO92
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AC0173
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0173
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-370 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC90246.1; PID:g15979466; GSPDB:GN00175
C;Genetics:
A;Gene: YPO1417

Query Match 100.0%; Score 29; DB 2; Length 370;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 266 TKPPR 270

RESULT 12
D84715
probable protein kinase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: D84715
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84715
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-375 <STO>
A;Cross-references: GB:AE002093; NID:g3201626; PIDN:AAC20735.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g31010
A;Map position: 2

Query Match 100.0%; Score 29; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 121 TKPPR 125

RESULT 13
S38238
hypothetical protein - Coxiella burnetii
C;Species: Coxiella burnetii
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
C;Accession: S38238
R;Thiele, D.; Willems, H.; Haas, M.; Krauss, H.
submitted to the EMBL Data Library, October 1993
A;Reference number: S38215
A;Accession: S38238
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-410 <THI>
A;Cross-references: EMBL:X75356; NID:g407370; PIDN:CAA53126.1; PID:g407394

Query Match 100.0%; Score 29; DB 2; Length 410;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              |||||
Db        102 TKPPR 106
```

RESULT 14
F96499
hypothetical protein T10P12.9 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: F96499
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96499
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-415 <STO>
A;Cross-references: GB:AE005173; NID:g5080765; PIDN:AAD39275.1; GSPDB:GN00141
C;Genetics:
A;Gene: T10P12.9
A;Map position: 1

Query Match 100.0%; Score 29; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              |||||
Db        309 TKPPR 313
```

RESULT 15
C84829
hypothetical protein At2g40420 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84829
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84829
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-415 <STO>
A;Cross-references: GB:AE002093; NID:g6598346; PIDN:AAB87575.2; GSPDB:GN00139
C;Genetics:
A;Gene: At2g40420
A;Map position: 2

Query Match 100.0%; Score 29; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              |||||
Db        169 TKPPR 173
```

RESULT 16
S61165
hypothetical protein YDR370c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein D9481.14
C;Species: Saccharomyces cerevisiae
C;Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 19-Apr-2002
C;Accession: S61165
R;Ding, H.
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of S. cerevisiae cosmid 9481.
A;Reference number: S61159
A;Accession: S61165
A;Molecule type: DNA
A;Residues: 1-442 <DIN>
A;Cross-references: EMBL:U28373; NID:g849184; PID:g849191; GSPDB:GN00004; MIPS:YDR370c
A;Experimental source: strain S288C (AB972)
C;Genetics:
A;Gene: MIPS:YDR370c
A;Cross-references: SGD:S0002778
A;Map position: 4R
C;Superfamily: Saccharomyces cerevisiae hypothetical protein YDR370c

Query Match 100.0%; Score 29; DB 2; Length 442;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              |||||
Db         81 TKPPR 85
```

RESULT 17
S73752
hypothetical protein A05_orf493 - Mycoplasma pneumoniae (strain ATCC 29342)
C;Species: Mycoplasma pneumoniae
A;Variety: ATCC 29342
C;Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 15-Sep-2000
C;Accession: S73752
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumonia
A;Reference number: S73327; MUID:97105885; PMID:8948633
A;Accession: S73752
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-493 <HIM>
A;Cross-references: EMBL:AE000041; GB:U00089; NID:g1674104; PIDN:AAB96074.1; PID:g16741
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C;Genetics:
A;Genetic code: SGC3
C;Superfamily: Mycoplasma pneumoniae hypothetical protein H08_orf445

Query Match 100.0%; Score 29; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              |||||
Db        462 TKPPR 466
```

RESULT 18
T21775
hypothetical protein F35E2.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21775
R;Lennard, N.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19471
A;Accession: T21775
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA

A;Residues: 1-575 <WIL>
A;Cross-references: EMBL:Z81528; PIDN:CAB04281.1; GSPDB:GN00019; CESP:F35E2.5
A;Experimental source: clone F35E2
C;Genetics:
A;Gene: CESP:F35E2.5
A;Map position: 1
A;Introns: 44/1; 109/2; 209/3; 250/1; 367/1; 413/1; 480/3; 508/3

Query Match 100.0%; Score 29; DB 2; Length 575;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 284 TKPPR 288

RESULT 19
S56027
hypothetical protein YJL083w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein J1002
C;Species: Saccharomyces cerevisiae
C;Date: 05-May-1995 #sequence_revision 08-Sep-1995 #text_change 19-Apr-2002
C;Accession: S56027; S56860; S57742
R;Miosga, T.; Schaaff-Gerstenschlaeger, I.; Chalwatzis, N.; Baur, A.; Boles, E.; Fournie
Yeast 11, 681-689, 1995
A;Title: Sequence analysis of a 33.1 kb fragment from the left arm of Saccharomyces cere
ter domain and a putative alpha-2-SCB-alpha-2 binding site.
A;Reference number: S56016; MUID:96093911; PMID:7483841
A;Accession: S56027
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-604 <MIO>
A;Cross-references: EMBL:X83502; NID:g929861; PIDN:CAA58487.1; PID:g929873
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1994
R;Miosga, T.; Schaaff-Gerstenschlaeger, I.; Baur, A.; Boles, E.; Chalwatzis, N.; Fournie
submitted to the Protein Sequence Database, September 1995
A;Reference number: S56855
A;Accession: S56860
A;Molecule type: DNA
A;Residues: 1-604 <MIW>
A;Cross-references: EMBL:Z49358; NID:g1008247; PIDN:CAA89376.1; PID:g1008248; MIPS:YJL08
R;Sor, F.J.
submitted to the EMBL Data Library, June 1995
A;Reference number: S57731
A;Accession: S57742
A;Molecule type: DNA
A;Residues: 1-604 <SOR>
A;Cross-references: EMBL:X88851; NID:g895892; PIDN:CAA61318.1; PID:g895904
C;Genetics:
A;Cross-references: SGD:S0003619
A;Map position: 10L

Query Match 100.0%; Score 29; DB 2; Length 604;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 151 TKPPR 155

RESULT 20
JC7576
transcription factor Elf-1, type 1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C;Accession: JC7576
R;Nishiyama, C.; Takahashi, K.; Nishiyama, M.; Okumura, K.; Ra, C.; Ohtake, Y.; Yokota,
Biosci. Biotechnol. Biochem. 64, 2601-2607, 2000
A;Title: Splice isoforms of transcription factor Elf-1 affecting its regulatory function
A;Reference number: JC7576; MUID: 21077473; PMID:11210123
A;Contents: Mast cell line, RBL-2H3
A;Accession: JC7576
A;Molecule type: mRNA
A;Residues: 1-615 <NIS>
A;Cross-references: DDBJ:AB030215
C;Comment: This protein, as a key transcription factor for immune-related genes, has th
C;Genetics:
A;Gene: Elf-1

Query Match 100.0%; Score 29; DB 2; Length 615;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 179 TKPPR 183

RESULT 21
A43361
Ets-related transcription factor Elf-1 - human
N;Alternate names: E74-like factor Elf-1
C;Species: Homo sapiens (man)
C;Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 07-May-1999
C;Accession: A43361; A42122
R;Leiden, J.M.; Wang, C.Y.; Petryniak, B.; Markovitz, D.M.; Nabel, G.J.; Thompson, C.B
J. Virol. 66, 5890-5897, 1992
A;Title: A novel Ets-related transcription factor, Elf-1, binds to human immunodeficie
A;Reference number: A43361; MUID:92407982; PMID:1527846
A;Accession: A43361
A;Molecule type: mRNA
A;Residues: 1-619 <LEI>
A;Cross-references: GB:M82882
R;Thompson, C.B.; Wang, C.Y.; Ho, I.C.; Bohjanen, P.R.; Petryniak, B.; June, C.H.; Mie
Mol. Cell. Biol. 12, 1043-1053, 1992
A;Title: cis-acting sequences required for inducible interleukin-2 enhancer function b
A;Reference number: A42122; MUID:92186836; PMID:1545787
A;Accession: A42122
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 204-282,'G',284-289 <THO>
A;Experimental source: T-cells
A;Note: sequence extracted from NCBI backbone (NCBIP:88288)
C;Genetics:
A;Gene: GDB:ELF1
A;Cross-references: GDB:131648
A;Map position: 1p36-1p36
C;Superfamily: ets DNA-binding domain homology
C;Keywords: DNA binding; transcription regulation
F;210-290/Domain: ets DNA-binding domain homology <ETS>

Query Match 100.0%; Score 29; DB 2; Length 619;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 180 TKPPR 184

RESULT 22
T22359
hypothetical protein F47G4.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T22359
R;White, S.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z19553
A;Accession: T22359
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-722 <WIL>
A;Cross-references: EMBL:Z99171; PIDN:CAB16313.1; GSPDB:GN00019; CESP:F47G4.4

A;Experimental source: clone F47G4
C;Genetics:
A;Gene: CESP:F47G4.4
A;Map position: 1
A;Introns: 5/1; 43/3; 129/3; 176/3; 226/1; 259/3; 319/3; 365/3; 465/2; 576/1; 631/1; 688

Query Match 100.0%; Score 29; DB 2; Length 722;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||||
Db 402 TKPPR 406

RESULT 23
T45684
hypothetical protein F14P22.230 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Mar-2000
C;Accession: T45684
R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23011
A;Accession: T45684
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-816 <DAN>
A;Cross-references: EMBL:AL137082
A;Experimental source: cultivar Columbia; BAC clone F14P22
C;Genetics:
A;Map position: 3
A;Introns: 83/3; 151/3; 209/3; 296/2; 336/2; 366/2; 398/2; 429/2; 469/3; 562/1; 603/2; 6
A;Note: F14P22.230
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

Query Match 100.0%; Score 29; DB 2; Length 816;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||||
Db 625 TKPPR 629

RESULT 24
A41275
DNA ligase (ATP) (EC 6.5.1.1) I - human
N;Alternate names: DNA joinase; DNA repair enzyme; polydeoxyribonucleotide synthase (ATP
C;Species: Homo sapiens (man)
C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 03-Feb-2003
C;Accession: A36048; A41275
R;Barnes, D.E.; Johnston, L.H.; Kodama, K.; Tomkinson, A.E.; Lasko, D.D.; Lindahl, T.
Proc. Natl. Acad. Sci. U.S.A. 87, 6679-6683, 1990
A;Title: Human DNA ligase I cDNA: cloning and functional expression in Saccharomyces cer
A;Reference number: A36048; MUID:90370849; PMID:2204063
A;Accession: A36048
A;Molecule type: mRNA
A;Residues: 1-919 <BAR>
A;Cross-references: GB:M36067; NID:g187142; PIDN:AAA59518.1; PID:g187143
R;Petrini, J.H.J.; Huwiler, K.G.; Weaver, D.T.
Proc. Natl. Acad. Sci. U.S.A. 88, 7615-7619, 1991
A;Title: A wild-type DNA ligase I gene is expressed in Bloom's syndrome cells.
A;Reference number: A41275; MUID:91352039; PMID:1881902
A;Accession: A41275
A;Molecule type: mRNA
A;Residues: 716-753 <PET>
C;Genetics:
A;Gene: GDB:LIG1
A;Cross-references: GDB:127274; OMIM:126391
A;Map position: 19q13.3-19q13.3
C;Keywords: DNA repair; ligase; phosphoprotein
F;568/Active site: Lys (covalent AMP-binding) #status predicted

Query Match 100.0%; Score 29; DB 2; Length 919;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||||
Db 218 TKPPR 222

RESULT 25
T51135
ligand-gated channel-like protein precursor [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 28-Jul-2000
C;Accession: T51135
R;Meyerhoff, O.; Hedrich, R.; Becker, D.
submitted to the EMBL Data Library, July 1999
A;Description: Characterization of ligand-gated channel-like proteins in higher plants
A;Reference number: Z25308
A;Accession: T51135
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-941 <MEY>
A;Cross-references: EMBL:AF167355; PIDN:AAD47833.1
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: GLUR3
A;Map position: 1

Query Match 100.0%; Score 29; DB 2; Length 941;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||||
Db 468 TKPPR 472

RESULT 26
D86186
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: D86186
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonsc
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D86186
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-962 <STO>
A;Cross-references: GB:AE005172; NID:g2388577; PIDN:AAB71458.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

Query Match 100.0%; Score 29; DB 2; Length 962;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||||
Db 468 TKPPR 472

RESULT 27
T51137
ionotropic glutamate receptor homolog GLR4 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 28-Jul-2000
C;Accession: T51137
R;Davenport, R.J.; Kiegle, E.A.; Tester, M.
submitted to the EMBL Data Library, September 1999
A;Description: Cloning of an ionotropic glutamate receptor homolog from Arabidopsis thal
A;Reference number: Z25310
A;Accession: T51137
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-976 <DAV>
A;Cross-references: EMBL:AF183932; PIDN:AAF01294.1
C;Genetics:
A;Gene: GLR4

Query Match 100.0%; Score 29; DB 2; Length 976;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 468 TKPPR 472

RESULT 28
PWBYH
potassium transport protein TRK1, high-affinity - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein J0693; protein YJL129c
C;Species: Saccharomyces cerevisiae
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 27-Oct-2003
C;Accession: S05849; S56910
R;Gaber, R.F.; Styles, C.A.; Fink, G.R.
Mol. Cell. Biol. 8, 2848-2859, 1988
A;Title: TRK1 encodes a plasma membrane protein required for high-affinity potassium tra
A;Reference number: S05849; MUID:88302204; PMID:3043197
A;Accession: S05849
A;Molecule type: DNA
A;Residues: 1-1235 <GAB>
A;Cross-references: EMBL:M21328; NID:g171803; PIDN:AAA34728.1; PID:g171804
R;Cziepluch, C.; Kordes, E.; Pujol, A.; Jauniaux, J.C.
submitted to the Protein Sequence Database, September 1995
A;Reference number: S56891
A;Accession: S56910
A;Molecule type: DNA
A;Residues: 1-1235 <CZI>
A;Cross-references: EMBL:Z49404; NID:g1008329; PIDN:CAA89424.1; PID:g1008330; GSPDB:GN00
C;Genetics:
A;Gene: SGD:TRK1; MIPS:YJL129c
A;Cross-references: SGD:S0003665; MIPS:YJL129c
A;Map position: 10L
C;Superfamily: potassium transport protein TRK1/TRK2
C;Keywords: ion transport; potassium transport; transmembrane protein
F;50-70/Domain: transmembrane #status predicted <TM01>
F;72-96/Domain: transmembrane #status predicted <TM02>
F;107-127/Domain: transmembrane #status predicted <TM03>
F;778-800/Domain: transmembrane #status predicted <TM04>
F;812-834/Domain: transmembrane #status predicted <TM05>
F;839-859/Domain: transmembrane #status predicted <TM06>
F;862-882/Domain: transmembrane #status predicted <TM07>
F;898-918/Domain: transmembrane #status predicted <TM08>
F;923-943/Domain: transmembrane #status predicted <TM09>
F;972-991/Domain: transmembrane #status predicted <TM10>
F;1079-1096/Domain: transmembrane #status predicted <TM11>
F;1111-1132/Domain: transmembrane #status predicted <TM12>
F;738/Binding site: ATP (Lys) #status predicted

Query Match 100.0%; Score 29; DB 1; Length 1235;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 438 TKPPR 442

RESULT 29
JU0466
potassium transport protein TRK1, high-affinity - yeast (Saccharomyces cerevisiae) (str
N;Alternate names: protein YJL129c
C;Species: Saccharomyces cerevisiae
A;Variety: Saccharomyces uvarum
C;Date: 13-Jul-1990 #sequence_revision 28-Oct-1994 #text_change 27-Oct-2003
C;Accession: JU0466
R;Anderson, J.A.; Best, L.A.; Gaber, R.F.
Gene 99, 39-46, 1991
A;Title: Structural and functional conservation between the high-affinity K+ transporte
A;Reference number: JU0466; MUID:91216443; PMID:2022322
A;Accession: JU0466
A;Molecule type: DNA
A;Residues: 1-1241 <AND>
A;Cross-references: GB:M57508; NID:g171640; PIDN:AAA34661.1; PID:g171641
A;Note: the source is designated as Saccharomyces uvarum
C;Genetics:
A;Gene: SGD:TRK1
A;Cross-references: SGD:S0003665; MIPS:YJL129c
A;Map position: 10L
C;Superfamily: potassium transport protein TRK1/TRK2
C;Keywords: ion transport; potassium transport; transmembrane protein
F;50-70/Domain: transmembrane #status predicted <TM01>
F;78-98/Domain: transmembrane #status predicted <TM02>
F;107-117/Domain: transmembrane #status predicted <TM03>
F;784-806/Domain: transmembrane #status predicted <TM04>
F;819-840/Domain: transmembrane #status predicted <TM05>
F;844-864/Domain: transmembrane #status predicted <TM06>
F;868-888/Domain: transmembrane #status predicted <TM07>
F;904-924/Domain: transmembrane #status predicted <TM08>
F;929-949/Domain: transmembrane #status predicted <TM09>
F;1084-1104/Domain: transmembrane #status predicted <TM10>
F;1117-1137/Domain: transmembrane #status predicted <TM11>

Query Match 100.0%; Score 29; DB 2; Length 1241;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 444 TKPPR 448

RESULT 30
G81008
hypothetical protein NMB2072 [imported] - Neisseria meningitidis (strain MC58 serogrou
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: G81008
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: G81008
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-47 <TET>
A;Cross-references: GB:AE002557; GB:AE002098; NID:g7227332; PIDN:AAF42391.1; PID:g7227
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB2072

Query Match 89.7%; Score 26; DB 2; Length 47;
Best Local Similarity 80.0%; Pred. No. 85;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              |:|||
Db         20 TRPPR 24
```

RESULT 31
AH2135
hypothetical protein asl2639 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AH2135
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2135
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-50 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB74338.1; PID:g17131732; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: asl2639

Query Match 89.7%; Score 26; DB 2; Length 50;
Best Local Similarity 80.0%; Pred. No. 91;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              ||||:
Db         20 TKPPK 24
```

RESULT 32
C81143
hypothetical protein NMB0921 [imported] - Neisseria meningitidis (strain MC58 serogroup
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: C81143
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: C81143
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-66 <TET>
A;Cross-references: GB:AE002443; GB:AE002098; NID:g7226149; PIDN:AAF41329.1; PID:g722615
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0921

Query Match 89.7%; Score 26; DB 2; Length 66;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              ||||:
Db         39 TKPPK 43
```

RESULT 33
T36774
hypothetical protein SCI28.04c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T36774
R;Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A;Reference number: Z21574
A;Accession: T36774
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-68 <SEE>
A;Cross-references: EMBL:AL096844; PIDN:CAB50878.1; GSPDB:GN00070; SCOEDB:SCI28.04c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCI28.04c

Query Match 89.7%; Score 26; DB 2; Length 68;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              |:|||
Db          7 TRPPR 11
```

RESULT 34
AE1652
hypothetical protein lin1758 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AE1652
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeck
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, I
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; I
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlan
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AE1652
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-75 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC96989.1; PID:g16414245; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin1758

Query Match 89.7%; Score 26; DB 2; Length 75;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              ||||:
Db         38 TKPPK 42
```

RESULT 35
AD0288
conserved hypothetical protein YPO2363 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AD0288
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.I
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AD0288
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-87 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC91168.1; PID:g15980360; GSPDB:GN00175
C;Genetics:
A;Gene: YPO2363

```
  Query Match             89.7%;  Score 26;  DB 2;  Length 87;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db         83 TRPPR 87


RESULT 36
T17560
hypothetical protein a70L - Chlorella virus PBCV-1
C;Species: Chlorella virus PBCV-1
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T17560
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T17560
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-88 <GRA>
A;Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96438.1
A;Experimental source: specific host Chlorella strain NC64A
C;Genetics:
A;Gene: a70L

  Query Match             89.7%;  Score 26;  DB 2;  Length 88;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db         51 TKPPK 55


RESULT 37
F72587
hypothetical protein APE1171 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: F72587
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: F72587
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-100 <KAW>
A;Cross-references: DDBJ:AP000061; NID:g5104821; PIDN:BAA80156.1; PID:d1043942; PID:g510
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1171

  Query Match             89.7%;  Score 26;  DB 2;  Length 100;
  Best Local Similarity   80.0%;  Pred. No. 1.8e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db         70 TRPPR 74


RESULT 38
JC2202
secretin precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 16-Jul-1999
C;Accession: JC2202; S34214
R;Lan, M.S.; Kajiyama, W.; Donadel, G.; Lu, J.; Notkins, A.L.
Biochem. Biophys. Res. Commun. 200, 1066-1071, 1994
A;Title: cDNA sequence and genomic organization of mouse secretin.
A;Reference number: JC2202; MUID:94234995; PMID:8179583
A;Accession: JC2202
A;Molecule type: mRNA
A;Residues: 1-133 <LAN>
A;Cross-references: EMBL:X73580; NID:g313710; PIDN:CAA51982.1; PID:g313711
C;Comment: This protein regulates the secretion of pancreatic juices and stimulates in
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; duplication; hormone; secretagogue
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-133/Product: prosecretin #status predicted <PRO>
F;32-58/Product: secretin #status predicted <MAT>
F;58/Modified site: amidated carboxyl end (Val) (amide in mature form from following g

  Query Match             89.7%;  Score 26;  DB 2;  Length 133;
  Best Local Similarity   80.0%;  Pred. No. 2.4e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        129 TRPPR 133


RESULT 39
B49530
vascular endothelial growth factor homolog A2R, 14.7K - Orf virus
C;Species: Orf virus
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C;Accession: B49530
R;Lyttle, D.J.; Fraser, K.M.; Fleming, S.B.; Mercer, A.A.; Robinson, A.J.
J. Virol. 68, 84-92, 1994
A;Title: Homologs of vascular endothelial growth factor are encoded by the poxvirus or
A;Reference number: A49530; MUID:94076465; PMID:8254780
A;Contents: NZ2
A;Accession: B49530
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-133 <LYT>
A;Cross-references: GB:S67520; NID:g456897; PIDN:AAB29220.1; PID:g456899
A;Note: sequence inconsistent with nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:141420, NCBIP:141425)

  Query Match             89.7%;  Score 26;  DB 2;  Length 133;
  Best Local Similarity   80.0%;  Pred. No. 2.4e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        126 TRPPR 130


RESULT 40
A29622
ribosomal protein S12, mitochondrial - fruit fly (Drosophila melanogaster)
N;Alternate names: technical knockout protein
C;Species: Drosophila melanogaster
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 12-Jun-2003
C;Accession: A29622
R;Royden, C.S.; Pirrotta, V.; Jan, L.Y.
Cell 51, 165-173, 1987
A;Title: The tko locus, site of a behavioral mutation in D. melanogaster, codes for a
A;Reference number: A29622; MUID:88027001; PMID:3117373
A;Accession: A29622
A;Molecule type: mRNA
A;Residues: 1-140 <ROY>
A;Cross-references: GB:M19494; NID:g158601; PIDN:AAA28935.1; PID:g158602
C;Genetics:
A;Gene: tko
A;Cross-references: FlyBase:FBgn0003714
C;Superfamily: ribosomal protein S12
C;Keywords: mitochondrion; protein biosynthesis; ribosome
F;117/Modified site: beta-methylthioaspartic acid (Asp) #status predicted
```

Query Match 89.7%; Score 26; DB 2; Length 140;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|:|||
Db 45 TRPPR 49

RESULT 41
T27059
hypothetical protein Y51A2A.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27059
R;McMurray, A.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z20304
A;Accession: T27059
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-140 <WIL>
A;Cross-references: EMBL:AL032635; PIDN:CAA21601.1; GSPDB:GN00023; CESP:Y51A2A.6
A;Experimental source: clone Y51A2A
C;Genetics:
A;Gene: CESP:Y51A2A.6
A;Map position: 5
A;Introns: 93/3; 129/1

Query Match 89.7%; Score 26; DB 2; Length 140;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|:|||
Db 68 TRPPR 72

RESULT 42
T49706
hypothetical protein B23L21.190 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49706
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49706
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-144 <SCH>
A;Cross-references: EMBL:AL356172; GSPDB:GN00116; NCSP:B23L21.190
A;Experimental source: BAC clone B23L21; strain OR74A
C;Genetics:
A;Gene: NCSP:B23L21.190
A;Map position: 6

Query Match 89.7%; Score 26; DB 2; Length 144;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|:|||
Db 130 TRPPR 134

RESULT 43
A87466
hypothetical protein CC1749 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: A87466
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kol
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87466
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-148 <STO>
A;Cross-references: GB:AE005673; NID:g13423171; PIDN:AAK23725.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC1749

Query Match 89.7%; Score 26; DB 2; Length 148;
Best Local Similarity 80.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|:|||
Db 82 TRPPR 86

RESULT 44
S75561
hypothetical protein sll0812 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S75561
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocyst
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75561
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-153 <KAN>
A;Cross-references: EMBL:D90911; GB:AB001339; NID:g1653083; PIDN:BAA18122.1; PID:g1653
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: Synechocystis hypothetical protein sll0812

Query Match 89.7%; Score 26; DB 2; Length 153;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||:
Db 15 TKPPK 19

RESULT 45
A69267
transposase homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C;Accession: A69267
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archa
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: A69267
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-154 <KLE>
A;Cross-references: GB:AE001097; GB:AE000782; NID:g2689420; PIDN:AAB91094.1; PID:g2650

```
  Query Match             89.7%;  Score 26;  DB 2;  Length 154;
  Best Local Similarity   80.0%;  Pred. No. 2.8e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db         24 TKPPK 28


RESULT 46
T49798
conserved hypothetical protein [imported] - Neurospora crassa (fragment)
N;Alternate names: protein B9J10.350
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49798
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49798
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-163 <SCH>
A;Cross-references: EMBL:AL356324; GSPDB:GN00116; NCSP:B9J10.350
A;Experimental source: BAC clone B9J10; strain OR74A
C;Genetics:
A;Gene: NCSP:B9J10.350
A;Map position: 6

  Query Match             89.7%;  Score 26;  DB 2;  Length 163;
  Best Local Similarity   80.0%;  Pred. No. 3e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db         45 TRPPR 49


RESULT 47
AB1974
hypothetical protein alr1341 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AB1974
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB1974
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-179 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB73298.1; PID:g17130688; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr1341

  Query Match             89.7%;  Score 26;  DB 2;  Length 179;
  Best Local Similarity   80.0%;  Pred. No. 3.3e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              ||||:
Db         72 TKPPK 76


RESULT 48
H83561
hypothetical protein PA0671 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: H83561
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; E
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83561
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-183 <STO>
A;Cross-references: GB:AE004502; GB:AE004091; NID:g9946547; PIDN:AAG04060.1; GSPDB:GN0
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA0671

  Query Match             89.7%;  Score 26;  DB 2;  Length 183;
  Best Local Similarity   80.0%;  Pred. No. 3.3e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        157 TRPPR 161


RESULT 49
S45627
acidic endoproteinase precursor - Myxococcus xanthus (strain DK101)
C;Species: Myxococcus xanthus
A;Variety: strain DK101
C;Date: 10-Dec-1994 #sequence_revision 26-Apr-1996 #text_change 08-Oct-1999
C;Accession: S45627; S62857
R;Lucas, N.; Mazaud-Aujard, C.; Bremaud, L.; Cenatiempo, Y.; Julien, R.
Eur. J. Biochem. 222, 247-254, 1994
A;Title: Protein purification, gene cloning and sequencing of an acidic endoprotease f
A;Reference number: S45627; MUID:94291618; PMID:8020464
A;Accession: S45627
A;Molecule type: DNA
A;Residues: 1-195 <LUC>
A;Cross-references: EMBL:X75892; NID:g516391; PIDN:CAA53499.1; PID:g516392
A;Experimental source: strain DK101
A;Accession: S62857
A;Molecule type: protein
A;Residues: 65-101 <LUW>
C;Genetics:
A;Gene: Maep
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-64/Domain: propeptide #status predicted <PRO>
F;65-195/Product: acidic endoproteinase #status experimental <MAT>

  Query Match             89.7%;  Score 26;  DB 2;  Length 195;
  Best Local Similarity   80.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |:|||
Db        148 TRPPR 152


RESULT 50
FOAD72
major core protein VII precursor - human adenovirus 2
C;Species: Mastadenovirus h2 (human adenovirus 2)
A;Note: host Homo sapiens (man)
C;Date: 18-Apr-1984 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C;Accession: C03837; A03836
R;Alestroem, P.; Akusjaervi, G.; Lager, M.; Yeh-kai, L.; Pettersson, U.
J. Biol. Chem. 259, 13980-13985, 1984
A;Title: Genes encoding the core proteins of adenovirus type 2.
A;Reference number: A03837; MUID:85054835; PMID:6094534
A;Accession: C03837
A;Molecule type: DNA
A;Residues: 1-198 <ALE>
```

A;Cross-references: GB:J01917; NID:g209811; PIDN:AAA92212.1; PID:g209828
R;Sung, M.T.; Cao, T.M.; Coleman, R.T.; Budelier, K.A.
Proc. Natl. Acad. Sci. U.S.A. 80, 2902-2906, 1983
A;Title: Gene and protein sequences of adenovirus protein VII, a hybrid basic chromosoma
A;Reference number: A03836; MUID:83221511; PMID:6574459
A;Accession: A03836
A;Molecule type: DNA
A;Residues: 1-111;113-198 <SUN>
A;Cross-references: GB:J01917
C;Genetics:
A;Map position: 43-45
C;Superfamily: adenovirus major core protein VII
C;Keywords: core protein; late protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-198/Product: major core protein VII #status predicted <MCP>

Query Match 89.7%; Score 26; DB 1; Length 198;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              |:|||
Db        193 TRPPR 197
```

Search completed: March 3, 2004, 12:19:05
Job time : 23 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 3, 2004, 12:16:13 ; Search time 23 Seconds
(without alignments)
11.223 Million cell updates/sec

Title: US-09-871-974-2
Perfect score: 29
Sequence: 1 TKPPR 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	5	1	US-08-171-737-2	Sequence 2, Appli
2	29	100.0	5	1	US-08-202-178-2	Sequence 2, Appli
3	29	100.0	5	1	US-08-299-636-37	Sequence 37, Appl
4	29	100.0	5	1	US-08-279-155-36	Sequence 36, Appl
5	29	100.0	5	1	US-08-713-484-2	Sequence 2, Appli
6	29	100.0	5	1	US-08-703-988A-36	Sequence 36, Appl
7	29	100.0	5	1	US-08-592-294-1	Sequence 1, Appli
8	29	100.0	5	1	US-08-454-859-1	Sequence 1, Appli
9	29	100.0	5	2	US-08-955-263-2	Sequence 2, Appli
10	29	100.0	5	2	US-08-612-842-36	Sequence 36, Appl
11	29	100.0	6	1	US-08-171-737-18	Sequence 18, Appl
12	29	100.0	6	1	US-08-299-636-38	Sequence 38, Appl
13	29	100.0	6	1	US-08-279-155-37	Sequence 37, Appl
14	29	100.0	6	1	US-08-703-988A-37	Sequence 37, Appl
15	29	100.0	6	2	US-08-612-842-37	Sequence 37, Appl
16	29	100.0	8	1	US-08-202-178-4	Sequence 4, Appli
17	29	100.0	8	1	US-08-713-484-4	Sequence 4, Appli
18	29	100.0	8	1	US-08-592-294-2	Sequence 2, Appli
19	29	100.0	8	2	US-08-955-263-4	Sequence 4, Appli
20	29	100.0	8	2	US-08-612-842-39	Sequence 39, Appl
21	29	100.0	9	1	US-08-171-737-9	Sequence 9, Appli
22	29	100.0	9	1	US-08-279-155-1	Sequence 1, Appli
23	29	100.0	9	1	US-08-279-155-38	Sequence 38, Appl
24	29	100.0	9	1	US-08-703-988A-1	Sequence 1, Appli
25	29	100.0	9	1	US-08-703-988A-38	Sequence 38, Appl
26	29	100.0	9	1	US-08-592-294-3	Sequence 3, Appli
27	29	100.0	9	1	US-08-454-859-2	Sequence 2, Appli
28	29	100.0	9	2	US-08-612-842-1	Sequence 1, Appli
29	29	100.0	9	2	US-08-612-842-38	Sequence 38, Appl
30	29	100.0	9	4	US-08-997-802-10	Sequence 10, Appl
31	29	100.0	9	4	US-08-997-802-11	Sequence 11, Appl
32	29	100.0	10	1	US-08-202-178-5	Sequence 5, Appli
33	29	100.0	10	1	US-08-713-484-5	Sequence 5, Appli
34	29	100.0	10	2	US-08-955-263-5	Sequence 5, Appli
35	29	100.0	11	1	US-08-202-178-7	Sequence 7, Appli
36	29	100.0	11	1	US-08-713-484-7	Sequence 7, Appli
37	29	100.0	11	2	US-08-955-263-7	Sequence 7, Appli
38	29	100.0	11	4	US-09-387-715-51	Sequence 51, Appl
39	29	100.0	11	4	US-09-387-715-52	Sequence 52, Appl
40	29	100.0	12	1	US-08-299-636-39	Sequence 39, Appl
41	29	100.0	60	4	US-09-381-546-13	Sequence 13, Appl
42	29	100.0	98	4	US-09-252-991A-25421	Sequence 25421, A
43	29	100.0	136	4	US-09-252-991A-21156	Sequence 21156, A
44	29	100.0	420	4	US-09-489-039A-8820	Sequence 8820, Ap
45	29	100.0	576	4	US-09-367-206-1	Sequence 1, Appli
46	29	100.0	576	4	US-09-367-206-21	Sequence 21, Appl
47	29	100.0	576	4	US-09-367-206-22	Sequence 22, Appl
48	29	100.0	576	4	US-09-367-206-23	Sequence 23, Appl
49	29	100.0	892	3	US-08-857-076-42	Sequence 42, Appl
50	29	100.0	1235	1	US-08-118-101A-2	Sequence 2, Appli
51	26	89.7	14	4	US-09-608-285A-44	Sequence 44, Appl
52	26	89.7	14	4	US-09-557-800C-44	Sequence 44, Appl
53	26	89.7	29	4	US-09-287-849-30	Sequence 30, Appl
54	26	89.7	44	4	US-09-645-470-14	Sequence 14, Appl
55	26	89.7	67	4	US-09-107-532A-3781	Sequence 3781, Ap
56	26	89.7	128	4	US-09-252-991A-28839	Sequence 28839, A
57	26	89.7	132	4	US-09-125-642C-15	Sequence 15, Appl
58	26	89.7	132	4	US-09-431-888-11	Sequence 11, Appl
59	26	89.7	133	4	US-09-431-888-2	Sequence 2, Appli
60	26	89.7	140	4	US-09-732-210-1700	Sequence 1700, Ap
61	26	89.7	147	4	US-09-252-991A-19713	Sequence 19713, A
62	26	89.7	178	4	US-09-252-991A-18672	Sequence 18672, A
63	26	89.7	211	4	US-09-252-991A-26873	Sequence 26873, A
64	26	89.7	211	4	US-09-252-991A-30029	Sequence 30029, A
65	26	89.7	220	4	US-09-252-991A-16739	Sequence 16739, A
66	26	89.7	234	4	US-09-252-991A-28621	Sequence 28621, A
67	26	89.7	245	4	US-09-252-991A-32366	Sequence 32366, A
68	26	89.7	253	4	US-09-252-991A-20329	Sequence 20329, A
69	26	89.7	263	1	US-08-565-386-9	Sequence 9, Appli
70	26	89.7	264	4	US-09-252-991A-22289	Sequence 22289, A
71	26	89.7	282	4	US-09-252-991A-19978	Sequence 19978, A
72	26	89.7	290	4	US-09-564-951B-4	Sequence 4, Appli
73	26	89.7	304	4	US-09-355-040-14	Sequence 14, Appl
74	26	89.7	307	4	US-09-355-040-13	Sequence 13, Appl
75	26	89.7	308	4	US-09-050-739-94	Sequence 94, Appl
76	26	89.7	309	4	US-09-252-991A-17744	Sequence 17744, A
77	26	89.7	313	4	US-09-540-236-2724	Sequence 2724, Ap
78	26	89.7	314	3	US-08-612-973-42	Sequence 42, Appl
79	26	89.7	314	3	US-08-927-597-42	Sequence 42, Appl
80	26	89.7	316	4	US-09-252-991A-30881	Sequence 30881, A
81	26	89.7	319	3	US-08-612-973-44	Sequence 44, Appl
82	26	89.7	319	3	US-08-927-597-44	Sequence 44, Appl
83	26	89.7	321	4	US-09-252-991A-24925	Sequence 24925, A
84	26	89.7	337	4	US-09-543-681A-5632	Sequence 5632, Ap
85	26	89.7	338	3	US-08-612-973-38	Sequence 38, Appl
86	26	89.7	338	3	US-08-927-597-38	Sequence 38, Appl
87	26	89.7	343	3	US-08-612-973-40	Sequence 40, Appl
88	26	89.7	343	3	US-08-927-597-40	Sequence 40, Appl
89	26	89.7	349	4	US-09-252-991A-30578	Sequence 30578, A
90	26	89.7	350	4	US-09-818-780-95	Sequence 95, Appl
91	26	89.7	359	4	US-09-269-137-8	Sequence 8, Appli
92	26	89.7	369	4	US-09-355-040-30	Sequence 30, Appl
93	26	89.7	375	4	US-09-540-236-2085	Sequence 2085, Ap
94	26	89.7	384	4	US-09-252-991A-18786	Sequence 18786, A
95	26	89.7	393	4	US-09-107-532A-3712	Sequence 3712, Ap
96	26	89.7	395	3	US-08-928-442-1	Sequence 1, Appli
97	26	89.7	428	1	US-08-570-157-5	Sequence 5, Appli
98	26	89.7	428	3	US-08-029-170-31	Sequence 31, Appl
99	26	89.7	428	4	US-09-076-510-5	Sequence 5, Appli
100	26	89.7	428	4	US-09-004-349-5	Sequence 5, Appli

101	26	89.7	442	1	US-08-220-151-22	Sequence 22, Appl
102	26	89.7	442	1	US-08-413-118-22	Sequence 22, Appl
103	26	89.7	442	3	US-08-473-446-22	Sequence 22, Appl
104	26	89.7	442	6	5470718-2	Patent No. 5470718
105	26	89.7	442	6	5470718-3	Patent No. 5470718
106	26	89.7	456	4	US-09-240-639-2	Sequence 2, Appli
107	26	89.7	463	3	US-08-612-973-46	Sequence 46, Appl
108	26	89.7	463	3	US-08-927-597-46	Sequence 46, Appl
109	26	89.7	466	1	US-07-923-739-2	Sequence 2, Appli
110	26	89.7	471	4	US-09-608-285A-60	Sequence 60, Appl
111	26	89.7	484	4	US-09-608-285A-27	Sequence 27, Appl
112	26	89.7	484	4	US-09-370-265-27	Sequence 27, Appl
113	26	89.7	484	4	US-09-557-800C-27	Sequence 27, Appl
114	26	89.7	484	4	US-09-252-991A-19511	Sequence 19511, A
115	26	89.7	484	4	US-09-370-625A-27	Sequence 27, Appl
116	26	89.7	490	3	US-08-612-973-36	Sequence 36, Appl
117	26	89.7	490	3	US-08-927-597-36	Sequence 36, Appl
118	26	89.7	492	4	US-09-107-532A-7062	Sequence 7062, Ap
119	26	89.7	496	4	US-09-252-991A-31407	Sequence 31407, A
120	26	89.7	524	4	US-09-252-991A-29707	Sequence 29707, A
121	26	89.7	666	4	US-09-050-739-70	Sequence 70, Appl
122	26	89.7	692	3	US-08-612-973-48	Sequence 48, Appl
123	26	89.7	692	3	US-08-927-597-48	Sequence 48, Appl
124	26	89.7	710	4	US-09-252-991A-20733	Sequence 20733, A
125	26	89.7	809	3	US-08-612-973-50	Sequence 50, Appl
126	26	89.7	809	3	US-08-927-597-50	Sequence 50, Appl
127	26	89.7	961	5	PCT-US93-11725-4	Sequence 4, Appli
128	26	89.7	977	4	US-09-252-991A-16655	Sequence 16655, A
129	26	89.7	988	4	US-09-252-991A-29699	Sequence 29699, A
130	26	89.7	1053	4	US-09-252-991A-26140	Sequence 26140, A
131	26	89.7	1385	2	US-08-687-399-7	Sequence 7, Appli
132	26	89.7	1968	1	US-07-745-206A-7	Sequence 7, Appli
133	26	89.7	1968	1	US-08-455-543A-45	Sequence 45, Appl
134	26	89.7	1968	2	US-08-223-305C-45	Sequence 45, Appl
135	26	89.7	1968	2	US-08-311-363-7	Sequence 7, Appli
136	26	89.7	2353	4	US-08-984-709A-50	Sequence 50, Appl
137	26	89.7	3033	1	US-07-925-695-8	Sequence 8, Appli
138	26	89.7	3033	1	US-07-925-695-9	Sequence 9, Appli
139	25	86.2	22	2	US-08-476-062A-12	Sequence 12, Appl
140	25	86.2	22	2	US-08-476-062A-16	Sequence 16, Appl
141	25	86.2	22	5	PCT-US96-01314-12	Sequence 12, Appl
142	25	86.2	22	5	PCT-US96-01314-16	Sequence 16, Appl
143	25	86.2	22	5	PCT-US96-01314-72	Sequence 72, Appl
144	25	86.2	36	4	US-09-050-516-43	Sequence 43, Appl
145	25	86.2	36	4	US-10-278-547-43	Sequence 43, Appl
146	25	86.2	67	4	US-09-252-991A-27687	Sequence 27687, A
147	25	86.2	91	4	US-09-489-039A-8900	Sequence 8900, Ap
148	25	86.2	103	1	US-08-376-296-10	Sequence 10, Appl
149	25	86.2	104	4	US-09-621-976-4724	Sequence 4724, Ap
150	25	86.2	111	4	US-09-149-476-673	Sequence 673, App

ALIGNMENTS

```
RESULT 1
US-08-171-737-2
; Sequence 2, Application US/08171737
; Patent No. 5480970
;  GENERAL INFORMATION:
;    APPLICANT:  Pollak, Alfred
;    APPLICANT:  Goodbody, Anne
;    TITLE OF INVENTION:  METAL CHELATORS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,737
;      FILING DATE:  22-DEC-1993
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/253/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-171-737-2

  Query Match           100.0%;  Score 29;  DB 1;  Length 5;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 2
US-08-202-178-2
; Sequence 2, Application US/08202178
; Patent No. 5569745
;  GENERAL INFORMATION:
;    APPLICANT:  Goodbody, Anne
;    APPLICANT:  Pollak, Alfred
;    TITLE OF INVENTION:  PEPTIDE-CHELATOR CONJUGATES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/202,178
;      FILING DATE:  25-FEB-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/258/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-202-178-2

  Query Match           100.0%;  Score 29;  DB 1;  Length 5;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5
```

RESULT 3
US-08-299-636-37
; Sequence 37, Application US/08299636
; Patent No. 5659041
; GENERAL INFORMATION:
; APPLICANT: POLLAK, Alfred
; APPLICANT: KIRBY, Robert A.
; APPLICANT: DUNN-DUFAULT, Robert
; TITLE OF INVENTION: HYDRAZINO-TYPE RADIONUCLIDE CHELATORS
; TITLE OF INVENTION: HAVING AN N3S CONFIGURATION
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,636
; FILING DATE: 02-SEP-1994
; CLASSIFICATION: 534
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/092,911
; FILING DATE: 18-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/262/ALLE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-299-636-37

Query Match 100.0%; Score 29; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5
```

RESULT 4
US-08-279-155-36
; Sequence 36, Application US/08279155
; Patent No. 5662885
; GENERAL INFORMATION:
; APPLICANT: POLLAK, Alfred
; APPLICANT: GOODBODY, Anne
; TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W., Suite 330 - G
; STREET: Street Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/279,155
; FILING DATE: 22-JUL-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MURRAY, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P8074-4005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-279-155-36

Query Match 100.0%; Score 29; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5
```

RESULT 5
US-08-713-484-2
; Sequence 2, Application US/08713484
; Patent No. 5679642
; GENERAL INFORMATION:
; APPLICANT: Goodbody, Anne
; APPLICANT: Pollak, Alfred
; TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,484
; FILING DATE: 13-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/202,178
; FILING DATE: 25-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/290/ALLE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399

```
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-713-484-2

  Query Match           100.0%;  Score 29;  DB 1;  Length 5;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 6
US-08-703-988A-36
; Sequence 36, Application US/08703988A
; Patent No. 5780006
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, Alfred
;    APPLICANT:  GOODBODY, Anne
;    TITLE OF INVENTION:  PEPTIDE DERIVED RADIONUCLIDE
;    TITLE OF INVENTION:  CHELATORS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM
;      ADDRESSEE:  LLP
;      STREET:  655 Fifteenth Street, N. W., Suite 330
;      STREET:  -
;      STREET:  G Street Lobby
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/703,988A
;      FILING DATE:  28-AUG-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/279,155
;      FILING DATE:  22-JUL-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURRAY, Robert B.
;      REGISTRATION NUMBER:  22,980
;      REFERENCE/DOCKET NUMBER:  P8074-6011
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-703-988A-36

  Query Match           100.0%;  Score 29;  DB 1;  Length 5;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 7
US-08-592-294-1
; Sequence 1, Application US/08592294
; Patent No. 5789555
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, ALFRED
;    APPLICANT:  DUNN-DUFAULT, ROBERT
;    TITLE OF INVENTION:  IMMOBILIZED LABELLING METHOD
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
;      STREET:  655 FIFTEENTH ST., N.W., SUITE 330-G STREET
;      STREET:  LOBBY
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,294
;      FILING DATE:
;      CLASSIFICATION:  534
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/CA94/00637
;      FILING DATE:  16-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/152,680
;      FILING DATE:  16-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURRAY, ROBERT B
;      REGISTRATION NUMBER:  22,980
;      REFERENCE/DOCKET NUMBER:  P8012-5002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-638-5000
;      TELEFAX:  202-638-4810
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-592-294-1

  Query Match           100.0%;  Score 29;  DB 1;  Length 5;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 8
US-08-454-859-1
; Sequence 1, Application US/08454859
; Patent No. 5804158
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, Alfred
;    TITLE OF INVENTION:  SEQUESTERED IMAGING AGENTS
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram
;      STREET:  655 Fifteenth Street, NW, Suite 330, G.
;      STREET:  Street Lobby
;      CITY:  Washington
```

```
;       STATE:  DC
;       COUNTRY:  USA
;       ZIP:  20005-5701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/454,859
;       FILING DATE:  31-MAY-1995
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Berman, Richard J.
;       REGISTRATION NUMBER:  39,107
;       REFERENCE/DOCKET NUMBER:  8074-5007
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 638-5000
;       TELEFAX:  (202) 638-4810
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-454-859-1

  Query Match             100.0%;  Score 29;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 9
US-08-955-263-2
; Sequence 2, Application US/08955263
; Patent No. 5866544
;  GENERAL INFORMATION:
;    APPLICANT:  Goodbody, Anne
;    APPLICANT:  Pollak, Alfred
;    TITLE OF INVENTION:  PEPTIDE-CHELATOR CONJUGATES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/955,263
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/713,484
;      FILING DATE:  13-SEP-1996
;      APPLICATION NUMBER:  US 08/202,178
;      FILING DATE:  25-FEB-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/290/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-955-263-2

  Query Match             100.0%;  Score 29;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 10
US-08-612-842-36
; Sequence 36, Application US/08612842
; Patent No. 5976495
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, ALFRED
;    APPLICANT:  GOODBODY, ANNE
;    TITLE OF INVENTION:  PEPTIDE DERIVED RADIONUCLIDE CHELATORS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM
;      STREET:  655 15TH STREET, NW, G STREET LOBBY, SUITE
;      STREET:  330
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/612,842
;      FILING DATE:  20-MAR-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERMAN, RICHARD J
;      REGISTRATION NUMBER:  39107
;      REFERENCE/DOCKET NUMBER:  8012-6002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202 638-5000
;      TELEFAX:  202 638-4810
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-612-842-36

  Query Match             100.0%;  Score 29;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 11
US-08-171-737-18
; Sequence 18, Application US/08171737
; Patent No. 5480970
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Pollak, Alfred
;    APPLICANT:  Goodbody, Anne
;    TITLE OF INVENTION:  METAL CHELATORS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,737
;      FILING DATE:  22-DEC-1993
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/253/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-171-737-18

  Query Match             100.0%;  Score 29;  DB 1;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          2 TKPPR 6


RESULT 12
US-08-299-636-38
; Sequence 38, Application US/08299636
; Patent No. 5659041
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, Alfred
;    APPLICANT:  KIRBY, Robert A.
;    APPLICANT:  DUNN-DUFAULT, Robert
;    TITLE OF INVENTION:  HYDRAZINO-TYPE RADIONUCLIDE CHELATORS
;    TITLE OF INVENTION:  HAVING AN N3S CONFIGURATION
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/299,636
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:  534
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/092,911
;      FILING DATE:  18-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/262/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-299-636-38

  Query Match             100.0%;  Score 29;  DB 1;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          2 TKPPR 6


RESULT 13
US-08-279-155-37
; Sequence 37, Application US/08279155
; Patent No. 5662885
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, Alfred
;    APPLICANT:  GOODBODY, Anne
;    TITLE OF INVENTION:  PEPTIDE DERIVED RADIONUCLIDE CHELATORS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
;      STREET:  655 Fifteenth Street, N. W., Suite 330 - G
;      STREET:  Street Lobby
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/279,155
;      FILING DATE:  22-JUL-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURRAY, Robert B.
;      REGISTRATION NUMBER:  22,980
;      REFERENCE/DOCKET NUMBER:  P8074-4005
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  6
;      OTHER INFORMATION:  /note= "Arg at position 6 has an OH
;      OTHER INFORMATION:  group."
US-08-279-155-37
```

```
  Query Match           100.0%;  Score 29;  DB 1;  Length 6;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          2 TKPPR 6


RESULT 14
US-08-703-988A-37
; Sequence 37, Application US/08703988A
; Patent No. 5780006
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, Alfred
;    APPLICANT:  GOODBODY, Anne
;    TITLE OF INVENTION:  PEPTIDE DERIVED RADIONUCLIDE
;    TITLE OF INVENTION:  CHELATORS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM
;      ADDRESSEE:  LLP
;      STREET:  655 Fifteenth Street, N. W., Suite 330
;      STREET:  -
;      STREET:  G Street Lobby
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/703,988A
;      FILING DATE:  28-AUG-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/279,155
;      FILING DATE:  22-JUL-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURRAY, Robert B.
;      REGISTRATION NUMBER:  22,980
;      REFERENCE/DOCKET NUMBER:  P8074-6011
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  6
;      OTHER INFORMATION:  /note= "Arg at position 6 has
;      OTHER INFORMATION:  an OH group."
US-08-703-988A-37

  Query Match           100.0%;  Score 29;  DB 1;  Length 6;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          2 TKPPR 6


RESULT 15
US-08-612-842-37
; Sequence 37, Application US/08612842
; Patent No. 5976495
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, ALFRED
;    APPLICANT:  GOODBODY, ANNE
;    TITLE OF INVENTION:  PEPTIDE DERIVED RADIONUCLIDE CHELATORS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM
;      STREET:  655 15TH STREET, NW, G STREET LOBBY, SUITE
;      STREET:  330
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/612,842
;      FILING DATE:  20-MAR-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERMAN, RICHARD J
;      REGISTRATION NUMBER:  39107
;      REFERENCE/DOCKET NUMBER:  8012-6002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202 638-5000
;      TELEFAX:  202 638-4810
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  6
;      OTHER INFORMATION:  /note= "Position 6 has an -OH
;      OTHER INFORMATION:  substituent."
US-08-612-842-37

  Query Match           100.0%;  Score 29;  DB 2;  Length 6;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          2 TKPPR 6


RESULT 16
US-08-202-178-4
; Sequence 4, Application US/08202178
; Patent No. 5569745
;  GENERAL INFORMATION:
;    APPLICANT:  Goodbody, Anne
;    APPLICANT:  Pollak, Alfred
;    TITLE OF INVENTION:  PEPTIDE-CHELATOR CONJUGATES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/202,178
;      FILING DATE:  25-FEB-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/258/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Ser substituted with
;      OTHER INFORMATION:  picolinic acid (Pic)."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  2
;      OTHER INFORMATION:  /note= "Cys substituted with
;      OTHER INFORMATION:  acetamidomethyl (Acm)."
US-08-202-178-4

  Query Match             100.0%;  Score 29;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          4 TKPPR 8


RESULT 17
US-08-713-484-4
; Sequence 4, Application US/08713484
; Patent No. 5679642
;  GENERAL INFORMATION:
;    APPLICANT:  Goodbody, Anne
;    APPLICANT:  Pollak, Alfred
;    TITLE OF INVENTION:  PEPTIDE-CHELATOR CONJUGATES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/713,484
;      FILING DATE:  13-SEP-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/202,178
;      FILING DATE:  25-FEB-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/290/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Ser substituted with
;      OTHER INFORMATION:  picolinic acid (Pic)."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  2
;      OTHER INFORMATION:  /note= "Cys substituted with
;      OTHER INFORMATION:  acetamidomethyl (Acm)."
US-08-713-484-4

  Query Match             100.0%;  Score 29;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          4 TKPPR 8


RESULT 18
US-08-592-294-2
; Sequence 2, Application US/08592294
; Patent No. 5789555
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, ALFRED
;    APPLICANT:  DUNN-DUFAULT, ROBERT
;    TITLE OF INVENTION:  IMMOBILIZED LABELLING METHOD
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
;      STREET:  655 FIFTEENTH ST., N.W., SUITE 330-G STREET
;      STREET:  LOBBY
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,294
;      FILING DATE:
;      CLASSIFICATION:  534
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/CA94/00637
;      FILING DATE:  16-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/152,680
;      FILING DATE:  16-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURRAY, ROBERT B
;      REGISTRATION NUMBER:  22,980
;      REFERENCE/DOCKET NUMBER:  P8012-5002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-638-5000
;      TELEFAX:  202-638-4810
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Picolinic acid is attached to Ser
;      OTHER INFORMATION:  of position 1."
US-08-592-294-2

  Query Match             100.0%;  Score 29;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          4 TKPPR 8


RESULT 19
US-08-955-263-4
; Sequence 4, Application US/08955263
; Patent No. 5866544
;  GENERAL INFORMATION:
;    APPLICANT:  Goodbody, Anne
;    APPLICANT:  Pollak, Alfred
;    TITLE OF INVENTION:  PEPTIDE-CHELATOR CONJUGATES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/955,263
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/713,484
;      FILING DATE:  13-SEP-1996
;      APPLICATION NUMBER:  US 08/202,178
;      FILING DATE:  25-FEB-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/290/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Ser substituted with
;      OTHER INFORMATION:  picolinic acid (Pic)."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  2
;      OTHER INFORMATION:  /note= "Cys substituted with
;      OTHER INFORMATION:  acetamidomethyl (Acm)."
US-08-955-263-4

  Query Match             100.0%;  Score 29;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          4 TKPPR 8


RESULT 20
US-08-612-842-39
; Sequence 39, Application US/08612842
; Patent No. 5976495
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, ALFRED
;    APPLICANT:  GOODBODY, ANNE
;    TITLE OF INVENTION:  PEPTIDE DERIVED RADIONUCLIDE CHELATORS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM
;      STREET:  655 15TH STREET, NW, G STREET LOBBY, SUITE
;      STREET:  330
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/612,842
;      FILING DATE:  20-MAR-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERMAN, RICHARD J
;      REGISTRATION NUMBER:  39107
;      REFERENCE/DOCKET NUMBER:  8012-6002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202 638-5000
;      TELEFAX:  202 638-4810
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Position 1 has a sarcosine
;      OTHER INFORMATION:  substituent."
US-08-612-842-39

  Query Match             100.0%;  Score 29;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          4 TKPPR 8


RESULT 21
US-08-171-737-9
; Sequence 9, Application US/08171737
; Patent No. 5480970
;  GENERAL INFORMATION:
;    APPLICANT:  Pollak, Alfred
;    APPLICANT:  Goodbody, Anne
;    TITLE OF INVENTION:  METAL CHELATORS
;    NUMBER OF SEQUENCES:  19
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,737
;      FILING DATE:  22-DEC-1993
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/253/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "X at position 1 is
;      OTHER INFORMATION:  picolinic acid (Pic)."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  3
;      OTHER INFORMATION:  /note= "Cys at position 3 is
;      OTHER INFORMATION:  substituted with acetamidomethyl (Acm)."
US-08-171-737-9

  Query Match             100.0%;  Score 29;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 22
US-08-279-155-1
; Sequence 1, Application US/08279155
; Patent No. 5662885
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, Alfred
;    APPLICANT:  GOODBODY, Anne
;    TITLE OF INVENTION:  PEPTIDE DERIVED RADIONUCLIDE CHELATORS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
;      STREET:  655 Fifteenth Street, N. W., Suite 330 - G
;      STREET:  Street Lobby
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/279,155
;      FILING DATE:  22-JUL-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURRAY, Robert B.
;      REGISTRATION NUMBER:  22,980
;      REFERENCE/DOCKET NUMBER:  P8074-4005
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Gly at position 1 has an
;      OTHER INFORMATION:  N,N-dimethyl group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  3
;      OTHER INFORMATION:  /note= "Cys at position 3 has an
;      OTHER INFORMATION:  Acm group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  9
;      OTHER INFORMATION:  /note= "Arg at position 10 is
;      OTHER INFORMATION:  unsubstituted or has an OH group."
US-08-279-155-1

  Query Match             100.0%;  Score 29;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 23
US-08-279-155-38
; Sequence 38, Application US/08279155
; Patent No. 5662885
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, Alfred
;    APPLICANT:  GOODBODY, Anne
;    TITLE OF INVENTION:  PEPTIDE DERIVED RADIONUCLIDE CHELATORS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
;      STREET:  655 Fifteenth Street, N. W., Suite 330 - G
;      STREET:  Street Lobby
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/279,155
;      FILING DATE:  22-JUL-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURRAY, Robert B.
;      REGISTRATION NUMBER:  22,980
;      REFERENCE/DOCKET NUMBER:  P8074-4005
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Gly at position 1 has an
;      OTHER INFORMATION:  N,N-dimethyl group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  3
;      OTHER INFORMATION:  /note= "Cys at position 3 has an
;      OTHER INFORMATION:  Acm group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  9
;      OTHER INFORMATION:  /note= "Arg at position 9 is
;      OTHER INFORMATION:  unsubstituted or has an OH group."
US-08-279-155-38

  Query Match           100.0%;  Score 29;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 24
US-08-703-988A-1
; Sequence 1, Application US/08703988A
; Patent No. 5780006
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, Alfred
;    APPLICANT:  GOODBODY, Anne
;    TITLE OF INVENTION:  PEPTIDE DERIVED RADIONUCLIDE
;    TITLE OF INVENTION:  CHELATORS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM
;      ADDRESSEE:  LLP
;      STREET:  655 Fifteenth Street, N. W., Suite 330
;      STREET:  -
;      STREET:  G Street Lobby
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/703,988A
;      FILING DATE:  28-AUG-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/279,155
;      FILING DATE:  22-JUL-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURRAY, Robert B.
;      REGISTRATION NUMBER:  22,980
;      REFERENCE/DOCKET NUMBER:  P8074-6011
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Gly at position 1 has
;      OTHER INFORMATION:  an N,N-dimethyl group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  3
;      OTHER INFORMATION:  /note= "Cys at position 3 has
;      OTHER INFORMATION:  an Acm group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  9
;      OTHER INFORMATION:  /note= "Arg at position 9 is
;      OTHER INFORMATION:  unsubstituted or has an OH group."
US-08-703-988A-1

  Query Match           100.0%;  Score 29;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 25
US-08-703-988A-38
; Sequence 38, Application US/08703988A
; Patent No. 5780006
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, Alfred
;    APPLICANT:  GOODBODY, Anne
;    TITLE OF INVENTION:  PEPTIDE DERIVED RADIONUCLIDE
;    TITLE OF INVENTION:  CHELATORS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM
;      ADDRESSEE:  LLP
;      STREET:  655 Fifteenth Street, N. W., Suite 330
;      STREET:  -
;      STREET:  G Street Lobby
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/703,988A
;      FILING DATE:  28-AUG-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/279,155
;      FILING DATE:  22-JUL-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURRAY, Robert B.
;      REGISTRATION NUMBER:  22,980
;      REFERENCE/DOCKET NUMBER:  P8074-6011
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Gly at position 1 has
;      OTHER INFORMATION:  an N,N-dimethyl group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  3
;      OTHER INFORMATION:  /note= "Cys at position 3 has
;      OTHER INFORMATION:  an Acm group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  9
;      OTHER INFORMATION:  /note= "Arg at position 9 is
;      OTHER INFORMATION:  unsubstituted or has an OH group."
US-08-703-988A-38

  Query Match             100.0%;  Score 29;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 26
US-08-592-294-3
; Sequence 3, Application US/08592294
; Patent No. 5789555
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, ALFRED
;    APPLICANT:  DUNN-DUFAULT, ROBERT
;    TITLE OF INVENTION:  IMMOBILIZED LABELLING METHOD
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
;      STREET:  655 FIFTEENTH ST., N.W., SUITE 330-G STREET
;      STREET:  LOBBY
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,294
;      FILING DATE:
;      CLASSIFICATION:  534
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/CA94/00637
;      FILING DATE:  16-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/152,680
;      FILING DATE:  16-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURRAY, ROBERT B
;      REGISTRATION NUMBER:  22,980
;      REFERENCE/DOCKET NUMBER:  P8012-5002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-638-5000
;      TELEFAX:  202-638-4810
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "N,N'-dimethyl is attached to Gly
;      OTHER INFORMATION:  of position 1."
US-08-592-294-3

  Query Match             100.0%;  Score 29;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 27
US-08-454-859-2
; Sequence 2, Application US/08454859
; Patent No. 5804158
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, Alfred
;    TITLE OF INVENTION:  SEQUESTERED IMAGING AGENTS
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram
;      STREET:  655 Fifteenth Street, NW, Suite 330, G.
;      STREET:  Street Lobby
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,859
;      FILING DATE:  31-MAY-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,107
;      REFERENCE/DOCKET NUMBER:  8074-5007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 638-5000
;      TELEFAX:  (202) 638-4810
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Position 1 has either a
;      OTHER INFORMATION:  dimethyl substituent or an N,N-dimethyl substituent."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  9
;      OTHER INFORMATION:  /note= "Position 9 has an -OH
```

```
  OTHER INFORMATION:  substituent."
-454-859-2

ry Match             100.0%;  Score 29;  DB 1;  Length 9;
t Local Similarity   100.0%;  Pred. No. 3e+05;
ches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

       1 TKPPR 5
         |||||
       5 TKPPR 9


LT 28
8-612-842-1
quence 1, Application US/08612842
tent No. 5976495
ENERAL INFORMATION:
 APPLICANT:  POLLAK, ALFRED
 APPLICANT:  GOODBODY, ANNE
 TITLE OF INVENTION:  PEPTIDE DERIVED RADIONUCLIDE CHELATORS
 NUMBER OF SEQUENCES:  39
 CORRESPONDENCE ADDRESS:
   ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM
   STREET:  655 15TH STREET, NW, G STREET LOBBY, SUITE
   STREET:  330
   CITY:  WASHINGTON
   STATE:  DC
   COUNTRY:  USA
   ZIP:  20005-5701
 COMPUTER READABLE FORM:
   MEDIUM TYPE:  Floppy disk
   COMPUTER:  IBM PC compatible
   OPERATING SYSTEM:  PC-DOS/MS-DOS
   SOFTWARE:  PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
   APPLICATION NUMBER:  US/08/612,842
   FILING DATE:  20-MAR-1996
   CLASSIFICATION:  424
 ATTORNEY/AGENT INFORMATION:
   NAME:  BERMAN, RICHARD J
   REGISTRATION NUMBER:  39107
   REFERENCE/DOCKET NUMBER:  8012-6002
 TELECOMMUNICATION INFORMATION:
   TELEPHONE:  202 638-5000
   TELEFAX:  202 638-4810
INFORMATION FOR SEQ ID NO:  1:
 SEQUENCE CHARACTERISTICS:
   LENGTH:  9 amino acids
   TYPE:  amino acid
   STRANDEDNESS:  single
   TOPOLOGY:  linear
 MOLECULE TYPE:  peptide
 FEATURE:
   NAME/KEY:  Modified-site
   LOCATION:  1
   OTHER INFORMATION:  /note= "Position 1 has either an
   OTHER INFORMATION:  N,N-dimethyl, an N,N-dibenzyl, an N,N-diethyl, or a sarcosine
   OTHER INFORMATION:  substituent."
 FEATURE:
   NAME/KEY:  Modified-site
   LOCATION:  9
   OTHER INFORMATION:  /note= "Position 9 is either
   OTHER INFORMATION:  unsubstituted or has an -OH substituent."
US-08-612-842-1

  Query Match             100.0%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 29
US-08-612-842-38
; Sequence 38, Application US/08612842
; Patent No. 5976495
;  GENERAL INFORMATION:
;    APPLICANT:  POLLAK, ALFRED
;    APPLICANT:  GOODBODY, ANNE
;    TITLE OF INVENTION:  PEPTIDE DERIVED RADIONUCLIDE CHELATORS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM
;      STREET:  655 15TH STREET, NW, G STREET LOBBY, SUITE
;      STREET:  330
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/612,842
;      FILING DATE:  20-MAR-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERMAN, RICHARD J
;      REGISTRATION NUMBER:  39107
;      REFERENCE/DOCKET NUMBER:  8012-6002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202 638-5000
;      TELEFAX:  202 638-4810
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Position 1 has an
;      OTHER INFORMATION:  N,N-dimethyl substituent."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  9
;      OTHER INFORMATION:  /note= "Position 9 is either
;      OTHER INFORMATION:  unsubstituted or has an -OH substituent."
US-08-612-842-38

  Query Match             100.0%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 30
US-08-997-802-10
; Sequence 10, Application US/08997802
; Patent No. 6334996
; GENERAL INFORMATION:
;  APPLICANT: WONG, Ernest
;  TITLE OF INVENTION: CHELATORS THAT PREDOMINATELY FORM A SINGLE
;  TITLE OF INVENTION: STEREOISOMERIC SPECIES UPON COORDINATION TO A METAL
;  TITLE OF INVENTION: CENTER
;  FILE REFERENCE: 8298-7019
;  CURRENT APPLICATION NUMBER: US/08/997,802
```

```
  CURRENT FILING DATE: 1997-12-24
  NUMBER OF SEQ ID NOS: 12
  SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 10
   LENGTH: 9
   TYPE: PRT
   ORGANISM: Artificial Sequence
   FEATURE:
   NAME/KEY: MOD_RES
   LOCATION: (1)
   OTHER INFORMATION: dimethylglycine
   FEATURE:
   NAME/KEY: MOD_RES
   LOCATION: (2)
   OTHER INFORMATION: L-t-butylglycine
   FEATURE:
   NAME/KEY: MOD_RES
   LOCATION: (3)
   OTHER INFORMATION: L-cysteine with an acetoamidomethyl protecting
   OTHER INFORMATION: group attached via the Sulfur atom.
   FEATURE:
   FEATURE:
   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
   OTHER INFORMATION: Peptide
S-08-997-802-10

  Query Match            100.0%;  Score 29;  DB 4;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

)y          1 TKPPR 5
              |||||
)b          5 TKPPR 9


RESULT 31
JS-08-997-802-11
; Sequence 11, Application US/08997802
; Patent No. 6334996
; GENERAL INFORMATION:
;  APPLICANT: WONG, Ernest
;  TITLE OF INVENTION: CHELATORS THAT PREDOMINATELY FORM A SINGLE
;  TITLE OF INVENTION: STEREOISOMERIC SPECIES UPON COORDINATION TO A METAL
;  TITLE OF INVENTION: CENTER
;  FILE REFERENCE: 8298-7019
;  CURRENT APPLICATION NUMBER: US/08/997,802
;  CURRENT FILING DATE: 1997-12-24
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (1)
;   OTHER INFORMATION: Dimethylglycine.
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (2)
;   OTHER INFORMATION: L-t-butyl glycine
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (3)
;   OTHER INFORMATION: L-cysteine.
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-08-997-802-11

  Query Match            100.0%;  Score 29;  DB 4;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 3e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 32
US-08-202-178-5
; Sequence 5, Application US/08202178
; Patent No. 5569745
;  GENERAL INFORMATION:
;    APPLICANT:  Goodbody, Anne
;    APPLICANT:  Pollak, Alfred
;    TITLE OF INVENTION:  PEPTIDE-CHELATOR CONJUGATES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/202,178
;      FILING DATE:  25-FEB-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/258/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Ser substituted with
;      OTHER INFORMATION:  picolinic acid (Pic)."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  2
;      OTHER INFORMATION:  /note= "Cys substituted with
;      OTHER INFORMATION:  acetamidomethyl (Acm)."
US-08-202-178-5

  Query Match            100.0%;  Score 29;  DB 1;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 16;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          6 TKPPR 10


RESULT 33
US-08-713-484-5
; Sequence 5, Application US/08713484
; Patent No. 5679642
;  GENERAL INFORMATION:
;    APPLICANT:  Goodbody, Anne
;    APPLICANT:  Pollak, Alfred
;    TITLE OF INVENTION:  PEPTIDE-CHELATOR CONJUGATES
```

```
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/713,484
;      FILING DATE:  13-SEP-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/202,178
;      FILING DATE:  25-FEB-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/290/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Ser substituted with
;      OTHER INFORMATION:  picolinic acid (Pic)."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  2
;      OTHER INFORMATION:  /note= "Cys substituted with
;      OTHER INFORMATION:  acetamidomethyl (Acm)."
US-08-713-484-5

  Query Match             100.0%;  Score 29;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          6 TKPPR 10


RESULT 34
US-08-955-263-5
; Sequence 5, Application US/08955263
; Patent No. 5866544
;  GENERAL INFORMATION:
;    APPLICANT:  Goodbody, Anne
;    APPLICANT:  Pollak, Alfred
;    TITLE OF INVENTION:  PEPTIDE-CHELATOR CONJUGATES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/955,263
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/713,484
;      FILING DATE:  13-SEP-1996
;      APPLICATION NUMBER:  US 08/202,178
;      FILING DATE:  25-FEB-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/290/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Ser substituted with
;      OTHER INFORMATION:  picolinic acid (Pic)."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  2
;      OTHER INFORMATION:  /note= "Cys substituted with
;      OTHER INFORMATION:  acetamidomethyl (Acm)."
US-08-955-263-5

  Query Match             100.0%;  Score 29;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          6 TKPPR 10


RESULT 35
US-08-202-178-7
; Sequence 7, Application US/08202178
; Patent No. 5569745
;  GENERAL INFORMATION:
;    APPLICANT:  Goodbody, Anne
;    APPLICANT:  Pollak, Alfred
;    TITLE OF INVENTION:  PEPTIDE-CHELATOR CONJUGATES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/202,178
;      FILING DATE:  25-FEB-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/258/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
```

```
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Ser is substituted with
;      OTHER INFORMATION:  benzoylmercaptoacetic acid (Bz-MA)."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  2
;      OTHER INFORMATION:  /note= "Cys is substituted with
;      OTHER INFORMATION:  acetamidomethyl (Acm)."
US-08-202-178-7

  Query Match             100.0%;  Score 29;  DB 1;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          7 TKPPR 11


RESULT 36
US-08-713-484-7
; Sequence 7, Application US/08713484
; Patent No. 5679642
;  GENERAL INFORMATION:
;    APPLICANT:  Goodbody, Anne
;    APPLICANT:  Pollak, Alfred
;    TITLE OF INVENTION:  PEPTIDE-CHELATOR CONJUGATES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/713,484
;      FILING DATE:  13-SEP-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/202,178
;      FILING DATE:  25-FEB-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/290/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Ser is substituted with
;      OTHER INFORMATION:  benzoylmercaptoacetic acid (Bz-MA)."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  2
;      OTHER INFORMATION:  /note= "Cys is substituted with
;      OTHER INFORMATION:  acetamidomethyl (Acm)."
US-08-713-484-7

  Query Match             100.0%;  Score 29;  DB 1;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          7 TKPPR 11


RESULT 37
US-08-955-263-7
; Sequence 7, Application US/08955263
; Patent No. 5866544
;  GENERAL INFORMATION:
;    APPLICANT:  Goodbody, Anne
;    APPLICANT:  Pollak, Alfred
;    TITLE OF INVENTION:  PEPTIDE-CHELATOR CONJUGATES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  Suite 500, 3000 K Street, N.W.
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/955,263
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/713,484
;      FILING DATE:  13-SEP-1996
;      APPLICATION NUMBER:  US 08/202,178
;      FILING DATE:  25-FEB-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/290/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "Ser is substituted with
;      OTHER INFORMATION:  benzoylmercaptoacetic acid (Bz-MA)."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  2
;      OTHER INFORMATION:  /note= "Cys is substituted with
;      OTHER INFORMATION:  acetamidomethyl (Acm)."
US-08-955-263-7

  Query Match             100.0%;  Score 29;  DB 2;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 17;
```

```
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          7 TKPPR 11


RESULT 38
US-09-387-715-51
; Sequence 51, Application US/09387715
; Patent No. 6551574
; GENERAL INFORMATION:
;  APPLICANT: Sharma, Shubh
;  TITLE OF INVENTION: Tuftsin Metallopeptides Analogs and Uses Thereof
;  FILE REFERENCE: 1173/1D794US1
;  CURRENT APPLICATION NUMBER: US/09/387,715
;  CURRENT FILING DATE:  1999-08-30
;  PRIOR APPLICATION NUMBER: PCT/US99/05693
;  PRIOR FILING DATE: 1999-03-18
;  NUMBER OF SEQ ID NOS: 55
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
;   LENGTH: 11
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: peptide
;   NAME/KEY: MOD_RES
;   LOCATION: 6
;   OTHER INFORMATION: Ahe
;   NAME/KEY: modified residue
;   LOCATION: 8,10
;   OTHER INFORMATION: D-amino acid
US-09-387-715-51

  Query Match             100.0%;  Score 29;  DB 4;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 39
US-09-387-715-52
; Sequence 52, Application US/09387715
; Patent No. 6551574
; GENERAL INFORMATION:
;  APPLICANT: Sharma, Shubh
;  TITLE OF INVENTION: Tuftsin Metallopeptides Analogs and Uses Thereof
;  FILE REFERENCE: 1173/1D794US1
;  CURRENT APPLICATION NUMBER: US/09/387,715
;  CURRENT FILING DATE:  1999-08-30
;  PRIOR APPLICATION NUMBER: PCT/US99/05693
;  PRIOR FILING DATE: 1999-03-18
;  NUMBER OF SEQ ID NOS: 55
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
;   LENGTH: 11
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: peptide
;   NAME/KEY: MOD_RES
;   LOCATION: 6
;   OTHER INFORMATION: Ahe
;   NAME/KEY: modified residue
;   LOCATION: 2,4
;   OTHER INFORMATION: D-amino acid
US-09-387-715-52

  Query Match             100.0%;  Score 29;  DB 4;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          7 TKPPR 11


RESULT 40
US-08-299-636-39
; Sequence 39, Application US/08299636
; Patent No. 5659041
; GENERAL INFORMATION:
;    APPLICANT:  POLLAK, Alfred
;    APPLICANT:  KIRBY, Robert A.
;    APPLICANT:  DUNN-DUFAULT, Robert
;    TITLE OF INVENTION:  HYDRAZINO-TYPE RADIONUCLIDE CHELATORS
;    TITLE OF INVENTION:  HAVING AN N3S CONFIGURATION
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/299,636
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:  534
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/092,911
;      FILING DATE:  18-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16777/262/ALLE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /product= "OTHER"
;      OTHER INFORMATION:  /note= "The Xaa at position 1 = sulphur"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  2
;      OTHER INFORMATION:  /product= "OTHER"
;      OTHER INFORMATION:  /note= "The Xaa at position 2 = Acm"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  3
;      OTHER INFORMATION:  /product= "OTHER"
;      OTHER INFORMATION:  /note= "The Xaa at position 3 = Mercaptoacetyl"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  5
;      OTHER INFORMATION:  /product= "OTHER"
```

```
;      OTHER INFORMATION:  /note= "The Xaa at position 5 = nitrogen"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  6
;      OTHER INFORMATION:  /product= "OTHER"
;      OTHER INFORMATION:  /note= "The Xaa at position 6 = methylhydrazino nicotinic aci
US-08-299-636-39

  Query Match             100.0%;  Score 29;  DB 1;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          8 TKPPR 12


RESULT 41
US-09-381-546-13
; Sequence 13, Application US/09381546
; Patent No. 6451976
; GENERAL INFORMATION:
;  APPLICANT:  Trigen Limited
;  TITLE OF INVENTION:  BI- OR MULTIFUNCTIONAL MOLECULES BASED ON A DENDROASPIN
;  TITLE OF INVENTION:  SCAFFOLD
;  FILE REFERENCE:  P41007WO
;  CURRENT APPLICATION NUMBER: US/09/381,546
;  CURRENT FILING DATE:  1999-09-20
;  PRIOR APPLICATION NUMBER:  PCT/GB98/00848
;  PRIOR FILING DATE:  1998-09-20
;  PRIOR APPLICATION NUMBER:  GB9705787.1
;  PRIOR FILING DATE:  1997-03-20
;  NUMBER OF SEQ ID NOS:  48
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 60
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: modified
;   OTHER INFORMATION: dendroaspin
US-09-381-546-13

  Query Match             100.0%;  Score 29;  DB 4;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 82;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          9 TKPPR 13


RESULT 42
US-09-252-991A-25421
; Sequence 25421, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 25421
;   LENGTH: 98
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25421

  Query Match             100.0%;  Score 29;  DB 4;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         72 TKPPR 76


RESULT 43
US-09-252-991A-21156
; Sequence 21156, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 21156
;   LENGTH: 136
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21156

  Query Match             100.0%;  Score 29;  DB 4;  Length 136;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         86 TKPPR 90


RESULT 44
US-09-489-039A-8820
; Sequence 8820, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 8820
;   LENGTH: 420
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8820

  Query Match             100.0%;  Score 29;  DB 4;  Length 420;
  Best Local Similarity   100.0%;  Pred. No. 4.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        270 TKPPR 274


RESULT 45
US-09-367-206-1
; Sequence 1, Application US/09367206
```

```
; Patent No. 6326482
; GENERAL INFORMATION:
;  APPLICANT:  Genentech, Inc.
;  TITLE OF INVENTION:  NSP Molecules
;  FILE REFERENCE:  P1223R1E
;  CURRENT APPLICATION NUMBER: US/09/367,206
;  CURRENT FILING DATE:  1999-08-09
;  PRIOR APPLICATION NUMBER:  PCT/US99/08847
;  PRIOR FILING DATE:  1999-04-23
;  PRIOR APPLICATION NUMBER:  US 60/082,767
;  PRIOR FILING DATE:  1998-04-23
;  PRIOR APPLICATION NUMBER:  US 60/113,296
;  PRIOR FILING DATE:  1998-12-22
;  NUMBER OF SEQ ID NOS:  35
; SEQ ID NO 1
;   LENGTH: 576
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-367-206-1

  Query Match             100.0%;  Score 29;  DB 4;  Length 576;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        211 TKPPR 215


RESULT 46
US-09-367-206-21
; Sequence 21, Application US/09367206
; Patent No. 6326482
; GENERAL INFORMATION:
;  APPLICANT:  Genentech, Inc.
;  TITLE OF INVENTION:  NSP Molecules
;  FILE REFERENCE:  P1223R1E
;  CURRENT APPLICATION NUMBER: US/09/367,206
;  CURRENT FILING DATE:  1999-08-09
;  PRIOR APPLICATION NUMBER:  PCT/US99/08847
;  PRIOR FILING DATE:  1999-04-23
;  PRIOR APPLICATION NUMBER:  US 60/082,767
;  PRIOR FILING DATE:  1998-04-23
;  PRIOR APPLICATION NUMBER:  US 60/113,296
;  PRIOR FILING DATE:  1998-12-22
;  NUMBER OF SEQ ID NOS:  35
; SEQ ID NO 21
;   LENGTH: 576
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Mutation of SEQ ID NO:1
US-09-367-206-21

  Query Match             100.0%;  Score 29;  DB 4;  Length 576;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        211 TKPPR 215


RESULT 47
US-09-367-206-22
; Sequence 22, Application US/09367206
; Patent No. 6326482
; GENERAL INFORMATION:
;  APPLICANT:  Genentech, Inc.
;  TITLE OF INVENTION:  NSP Molecules
;  FILE REFERENCE:  P1223R1E
;  CURRENT APPLICATION NUMBER: US/09/367,206
;  CURRENT FILING DATE:  1999-08-09
;  PRIOR APPLICATION NUMBER:  PCT/US99/08847
;  PRIOR FILING DATE:  1999-04-23
;  PRIOR APPLICATION NUMBER:  US 60/082,767
;  PRIOR FILING DATE:  1998-04-23
;  PRIOR APPLICATION NUMBER:  US 60/113,296
;  PRIOR FILING DATE:  1998-12-22
;  NUMBER OF SEQ ID NOS:  35
; SEQ ID NO 22
;   LENGTH: 576
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Mutation of SEQ ID NO:1
US-09-367-206-22

  Query Match             100.0%;  Score 29;  DB 4;  Length 576;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        211 TKPPR 215


RESULT 48
US-09-367-206-23
; Sequence 23, Application US/09367206
; Patent No. 6326482
; GENERAL INFORMATION:
;  APPLICANT:  Genentech, Inc.
;  TITLE OF INVENTION:  NSP Molecules
;  FILE REFERENCE:  P1223R1E
;  CURRENT APPLICATION NUMBER: US/09/367,206
;  CURRENT FILING DATE:  1999-08-09
;  PRIOR APPLICATION NUMBER:  PCT/US99/08847
;  PRIOR FILING DATE:  1999-04-23
;  PRIOR APPLICATION NUMBER:  US 60/082,767
;  PRIOR FILING DATE:  1998-04-23
;  PRIOR APPLICATION NUMBER:  US 60/113,296
;  PRIOR FILING DATE:  1998-12-22
;  NUMBER OF SEQ ID NOS:  35
; SEQ ID NO 23
;   LENGTH: 576
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Mutation of SEQ ID NO:1
US-09-367-206-23

  Query Match             100.0%;  Score 29;  DB 4;  Length 576;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        211 TKPPR 215


RESULT 49
US-08-857-076-42
; Sequence 42, Application US/08857076C
; Patent No. 6225120
; GENERAL INFORMATION:
;  APPLICANT: Ruvkun, Gary
;  APPLICANT: Kimura, Koutarou
;  APPLICANT: Patterson, Garth
;  APPLICANT: Ogg, Scott
;  APPLICANT: Paradis, Suzanne
;  APPLICANT: Tissenbaum, Heidi
;  APPLICANT: Morris, Jason
;  APPLICANT: Koweek, Allison
;  TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
;  TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
```

```
;  FILE REFERENCE: 00786/351001
;  CURRENT APPLICATION NUMBER: US/08/857,076C
;  CURRENT FILING DATE: 1997-05-15
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 42
;   LENGTH: 892
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
US-08-857-076-42

  Query Match             100.0%;  Score 29;  DB 3;  Length 892;
  Best Local Similarity   100.0%;  Pred. No. 9.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         81 TKPPR 85


RESULT 50
US-08-118-101A-2
; Sequence 2, Application US/08118101A
; Patent No. 5620892
;  GENERAL INFORMATION:
;    APPLICANT:  Kurtz, Stephen E.
;    APPLICANT:  Knickerbocker, Aron M.
;    APPLICANT:  McCullough, John R.
;    TITLE OF INVENTION:  A STRAIN OF SACCHAROMYCES CEREVISIAE
;    TITLE OF INVENTION:  EXPRESSING THE GENE ENCODING POTASSIUM TRANSPORTER MINK
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burton Rodney
;      STREET:  P.O. Box 4000
;      CITY:  Princeton
;      STATE:  New Jersey
;      COUNTRY:  U.S.A.
;      ZIP:  08543-4000
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/118,101A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaul, Timothy J.
;      REGISTRATION NUMBER:  33,111
;      REFERENCE/DOCKET NUMBER:  DC27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 252-5901
;      TELEFAX:  (609) 252-4526
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1235 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-118-101A-2

  Query Match             100.0%;  Score 29;  DB 1;  Length 1235;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        438 TKPPR 442


Search completed: March  3, 2004, 12:19:39
Job time : 24 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 3, 2004, 12:15:03 ; Search time 39 Seconds
(without alignments)
40.451 Million cell updates/sec

Title: US-09-871-974-2
Perfect score: 29
Sequence: 1 TKPPR 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	73	5	O17534	O17534 caenorhabdi
2	29	100.0	106	6	Q29253	Q29253 sus scrofa
3	29	100.0	115	12	Q98676	Q98676 simian cyto
4	29	100.0	142	16	Q9PGQ8	Q9pgq8 xylella fas
5	29	100.0	142	16	Q87EU9	Q87eu9 xylella fas
6	29	100.0	143	13	Q9I996	Q9i996 acipenser b
7	29	100.0	169	12	Q7T922	Q7t922 human adeno
8	29	100.0	190	16	Q8F4V4	Q8f4v4 leptospira
9	29	100.0	191	10	Q9M250	Q9m250 arabidopsis
10	29	100.0	198	12	Q98674	Q98674 simian cyto
11	29	100.0	212	10	Q8H435	Q8h435 oryza sativ
12	29	100.0	222	16	Q8U6C3	Q8u6c3 agrobacteri
13	29	100.0	226	5	O45855	O45855 caenorhabdi
14	29	100.0	227	11	Q810F4	Q810f4 rattus norv
15	29	100.0	227	11	Q7TML1	Q7tml1 mus musculu
16	29	100.0	256	10	Q8SA84	Q8sa84 zea mays (m
17	29	100.0	260	11	O35455	O35455 mus musculu
18	29	100.0	278	16	P74248	P74248 synechocyst
19	29	100.0	283	5	Q24859	Q24859 entamoeba h
20	29	100.0	316	4	O60440	O60440 homo sapien
21	29	100.0	325	10	Q9FLJ1	Q9flj1 arabidopsis
22	29	100.0	326	10	Q9FL70	Q9fl70 arabidopsis
23	29	100.0	334	11	Q8CAP5	Q8cap5 mus musculu
24	29	100.0	350	16	Q8UKK5	Q8ukk5 agrobacteri
25	29	100.0	351	5	Q23287	Q23287 caenorhabdi
26	29	100.0	364	4	O60439	O60439 homo sapien
27	29	100.0	370	16	Q8ZG89	Q8zg89 yersinia pe
28	29	100.0	375	10	O80865	O80865 arabidopsis
29	29	100.0	394	16	Q82NT5	Q82nt5 streptomyce
30	29	100.0	405	16	Q89JB7	Q89jb7 bradyrhizob
31	29	100.0	406	16	Q83A13	Q83a13 coxiella bu
32	29	100.0	409	11	Q9D9R4	Q9d9r4 mus musculu
33	29	100.0	409	11	Q8BHB7	Q8bhb7 mus musculu
34	29	100.0	410	2	Q45945	Q45945 coxiella bu
35	29	100.0	410	2	O52878	O52878 coxiella bu
36	29	100.0	415	10	Q9XIF9	Q9xif9 arabidopsis
37	29	100.0	415	10	Q8S8E1	Q8s8e1 arabidopsis
38	29	100.0	415	17	Q96Z07	Q96z07 sulfolobus
39	29	100.0	429	10	Q944S3	Q944s3 arabidopsis
40	29	100.0	432	11	Q9D2W5	Q9d2w5 mus musculu
41	29	100.0	442	3	Q06349	Q06349 saccharomyc
42	29	100.0	458	5	Q9BLQ2	Q9blq2 meloidogyne
43	29	100.0	460	10	Q94LM4	Q94lm4 oryza sativ
44	29	100.0	491	16	Q828Y7	Q828y7 streptomyce
45	29	100.0	499	16	Q82GE0	Q82ge0 streptomyce
46	29	100.0	505	16	Q7UT23	Q7ut23 rhodopirell
47	29	100.0	514	11	Q9D318	Q9d318 mus musculu
48	29	100.0	524	11	Q9EQY1	Q9eqy1 rattus norv
49	29	100.0	547	5	Q9VQ69	Q9vq69 drosophila
50	29	100.0	574	4	Q96HX2	Q96hx2 homo sapien
51	29	100.0	575	5	O62223	O62223 caenorhabdi
52	29	100.0	576	4	Q9Y2X4	Q9y2x4 homo sapien
53	29	100.0	576	4	Q9BRG2	Q9brg2 homo sapien
54	29	100.0	579	11	Q9EQY2	Q9eqy2 rattus norv
55	29	100.0	585	12	Q7TFT8	Q7tft8 rhesus cyto
56	29	100.0	615	11	Q9EQY3	Q9eqy3 rattus norv
57	29	100.0	619	4	Q8N6F6	Q8n6f6 homo sapien
58	29	100.0	669	10	Q8GXJ4	Q8gxj4 arabidopsis
59	29	100.0	678	16	Q8PKG1	Q8pkg1 xanthomonas
60	29	100.0	722	5	Q9XTS1	Q9xts1 caenorhabdi
61	29	100.0	776	11	Q8BQI6	Q8bqi6 mus musculu
62	29	100.0	808	2	O70021	O70021 saccharopol
63	29	100.0	809	10	Q8L737	Q8l737 arabidopsis
64	29	100.0	809	10	Q94AB2	Q94ab2 arabidopsis
65	29	100.0	816	10	Q9M2F3	Q9m2f3 arabidopsis
66	29	100.0	832	3	Q871X8	Q871x8 neurospora
67	29	100.0	854	5	Q9U0T9	Q9u0t9 leishmania
68	29	100.0	866	13	Q8JHU7	Q8jhu7 brachydanio
69	29	100.0	892	5	Q95QI7	Q95qi7 caenorhabdi
70	29	100.0	906	4	Q8N4W1	Q8n4w1 homo sapien
71	29	100.0	941	10	Q9SWD9	Q9swd9 arabidopsis
72	29	100.0	959	10	Q8LGM9	Q8lgm9 arabidopsis
73	29	100.0	962	10	O23048	O23048 arabidopsis
74	29	100.0	988	5	Q9N8K2	Q9n8k2 trypanosoma
75	29	100.0	1003	4	Q9HCH0	Q9hch0 homo sapien
76	29	100.0	1003	16	Q98GJ5	Q98gj5 rhizobium l
77	29	100.0	1022	5	Q27779	Q27779 schistosoma
78	29	100.0	1047	4	O60284	O60284 homo sapien
79	29	100.0	1067	5	Q7YU69	Q7yu69 drosophila
80	29	100.0	1413	5	Q9VK08	Q9vk08 drosophila
81	29	100.0	1430	11	Q8VHK2	Q8vhk2 rattus norv
82	29	100.0	1491	5	Q86BT9	Q86bt9 giardia lam
83	29	100.0	1667	5	Q9VHT6	Q9vht6 drosophila
84	29	100.0	1706	5	Q9VAF5	Q9vaf5 drosophila
85	26	89.7	33	2	Q7WRI5	Q7wri5 escherichia
86	26	89.7	42	16	Q8KE21	Q8ke21 chlorobium
87	26	89.7	47	16	Q9JXF4	Q9jxf4 neisseria m
88	26	89.7	50	16	Q8YTS4	Q8yts4 anabaena sp
89	26	89.7	65	2	Q7WRH8	Q7wrh8 escherichia

90	26	89.7	66	16	Q9JZS0	Q9jzs0 neisseria m
91	26	89.7	68	16	Q9S279	Q9s279 streptomyce
92	26	89.7	75	16	Q92B10	Q92b10 listeria in
93	26	89.7	81	17	Q8ZWE1	Q8zwe1 pyrobaculum
94	26	89.7	86	6	Q9GMF3	Q9gmf3 bos taurus
95	26	89.7	87	5	Q8MKP6	Q8mkp6 drosophila
96	26	89.7	87	16	Q8ZE23	Q8ze23 yersinia pe
97	26	89.7	88	12	Q89405	Q89405 paramecium
98	26	89.7	89	10	Q7XNQ7	Q7xnq7 oryza sativ
99	26	89.7	100	17	Q9YCU1	Q9ycu1 aeropyrum p
100	26	89.7	101	12	Q9J1A0	Q9j1a0 suid herpes
101	26	89.7	102	12	Q806X6	Q806x6 hepatitis c
102	26	89.7	105	16	Q88KU2	Q88ku2 pseudomonas
103	26	89.7	110	10	Q8S2F3	Q8s2f3 oryza sativ
104	26	89.7	117	16	Q7UUS8	Q7uus8 rhodopirell
105	26	89.7	123	2	Q7WX19	Q7wx19 alcaligenes
106	26	89.7	127	16	Q8DBG4	Q8dbg4 vibrio vuln
107	26	89.7	132	12	Q9YMF3	Q9ymf3 orf virus.
108	26	89.7	139	11	Q80ZS9	Q80zs9 mus musculu
109	26	89.7	139	16	Q8KB15	Q8kb15 chlorobium
110	26	89.7	140	5	Q9U1J6	Q9u1j6 drosophila
111	26	89.7	140	5	Q9XWP9	Q9xwp9 caenorhabdi
112	26	89.7	141	10	Q8S3B2	Q8s3b2 physcomitre
113	26	89.7	147	10	Q7X6P1	Q7x6p1 oryza sativ
114	26	89.7	148	16	Q9A7H4	Q9a7h4 caulobacter
115	26	89.7	149	17	Q8TV08	Q8tv08 methanopyru
116	26	89.7	153	15	Q06322	Q06322 murine leuk
117	26	89.7	153	16	P74047	P74047 synechocyst
118	26	89.7	154	5	Q9NEH6	Q9neh6 drosophila
119	26	89.7	154	17	O30100	O30100 archaeoglob
120	26	89.7	156	16	Q8R8V4	Q8r8v4 thermoanaer
121	26	89.7	157	3	Q871H3	Q871h3 neurospora
122	26	89.7	163	5	Q8MZM8	Q8mzm8 taenia soli
123	26	89.7	167	4	Q8WTX0	Q8wtx0 homo sapien
124	26	89.7	171	16	Q82N66	Q82n66 streptomyce
125	26	89.7	172	4	Q9NWD7	Q9nwd7 homo sapien
126	26	89.7	173	12	Q88184	Q88184 san miguel
127	26	89.7	179	16	Q8YX74	Q8yx74 anabaena sp
128	26	89.7	183	10	Q7XRQ8	Q7xrq8 oryza sativ
129	26	89.7	183	11	Q9CWA7	Q9cwa7 mus musculu
130	26	89.7	183	16	Q9I5Q0	Q9i5q0 pseudomonas
131	26	89.7	185	5	O61697	O61697 manduca sex
132	26	89.7	189	10	Q9ARV1	Q9arv1 oryza sativ
133	26	89.7	192	12	Q8B8T7	Q8b8t7 human adeno
134	26	89.7	192	12	Q7T940	Q7t940 human adeno
135	26	89.7	192	12	Q7T8D1	Q7t8d1 human adeno
136	26	89.7	195	2	Q50902	Q50902 myxococcus
137	26	89.7	196	16	Q87XW1	Q87xw1 pseudomonas
138	26	89.7	198	5	Q7YVI3	Q7yvi3 trypanosoma
139	26	89.7	201	16	Q82SQ3	Q82sq3 nitrosomona
140	26	89.7	213	2	Q7X040	Q7x040 gamma-prote
141	26	89.7	213	12	Q8QS79	Q8qs79 chimpanzee
142	26	89.7	218	2	Q844P8	Q844p8 gamma-prote
143	26	89.7	218	2	Q844P7	Q844p7 gamma-prote
144	26	89.7	218	2	Q844N6	Q844n6 gamma-prote
145	26	89.7	220	12	O92486	O92486 bombyx mori
146	26	89.7	230	4	Q9UFE7	Q9ufe7 homo sapien
147	26	89.7	230	17	Q8ZV71	Q8zv71 pyrobaculum
148	26	89.7	231	10	Q7XVZ1	Q7xvz1 oryza sativ
149	26	89.7	231	12	Q8B6S1	Q8b6s1 flanders vi
150	26	89.7	232	17	Q9YA94	Q9ya94 aeropyrum p

ALIGNMENTS

```
RESULT 1
O17534
ID   O17534      PRELIMINARY;      PRT;    73 AA.
AC   O17534;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   DAF-3 (Fragment).
GN   DAF-3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RA   Patterson G., Koweek A., Wong A., Liu Y., Ruvkun G.;
RT   "The DAF-3 Smad protein antagonizes DAF-7 TGF-betareceptor signalling
RT   in the C. elegans dauer pathway.";
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF005207; AAB61750.1; -.
FT   NON_TER       1      1
FT   NON_TER      73     73
SQ   SEQUENCE   73 AA;  8061 MW;  7D854125D687F0EE CRC64;

  Query Match             100.0%;  Score 29;  DB 5;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 62;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         15 TKPPR 19


RESULT 2
Q29253
ID   Q29253      PRELIMINARY;      PRT;   106 AA.
AC   Q29253;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Cytochrome C oxidase polypeptide III (Fragment).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Small intestine;
RX   MEDLINE=96327607; PubMed=8672129;
RA   Winteroe A.K., Fredholm M., Davies W.;
RT   "Evaluation and characterization of a porcine small intestine cDNA
RT   library.";
RL   Mamm. Genome 7:509-517(1996).
DR   EMBL; F14871; CAA23306.1; -.
FT   NON_TER       1      1
FT   NON_TER     106    106
SQ   SEQUENCE   106 AA;  12552 MW;  4BF91EA407F2EF73 CRC64;

  Query Match             100.0%;  Score 29;  DB 6;  Length 106;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         48 TKPPR 52


RESULT 3
Q98676
ID   Q98676      PRELIMINARY;      PRT;   115 AA.
AC   Q98676;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   BHLF3 (Fragment).
OS   Simian cytomegalovirus.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Betaherpesvirinae; Cytomegalovirus.
OX   NCBI_TaxID=10364;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=11725047;
RA   Chang Y., Jeang K., Lietman T., Hayward G.S.;
RT   "Structural Organization of the Spliced Immediate-Early Gene Complex
RT   that Encodes the Major Acidic Nuclear (IE1) and Transactivator (IE2)
RT   Proteins of African Green Monkey Cytomegalovirus.";
RL   J. Biomed. Sci. 2:105-130(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90080130; PubMed=2152815;
RA   Chang Y.N., Crawford S., Stall J., Rawlins D.R., Jeang K.T.,
RA   Hayward G.S.;
RT   "The palindromic series I repeats in the simian cytomegalovirus major
RT   immediate-early promoter behave as both strong basal enhancers and
RT   cyclic AMP response elements.";
RL   J. Virol. 64:264-277(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87198884; PubMed=3033283;
RA   Jeang K.T., Rawlins D.R., Rosenfeld P.J., Shero J.H., Kelly T.J.,
RA   Hayward G.S.;
RT   "Multiple tandemly repeated binding sites for cellular nuclear factor
RT   1 that surround the major immediate-early promoters of simian and
RT   human cytomegalovirus.";
RL   J. Virol. 61:1559-1570(1987).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93100836; PubMed=8380090;
RA   Chang Y.N., Jeang K.T., Chiou C.J., Chan Y.J., Pizzorno M.,
RA   Hayward G.S.;
RT   "Identification of a large bent DNA domain and binding sites for serum
RT   response factor adjacent to the NFI repeat cluster and enhancer region
RT   in the major IE94 promoter from simian cytomegalovirus.";
RL   J. Virol. 67:516-529(1993).
DR   EMBL; U38308; AAB16875.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   115 AA;  12965 MW;  C4251671FBD87AF2 CRC64;

  Query Match             100.0%;  Score 29;  DB 12;  Length 115;
  Best Local Similarity   100.0%;  Pred. No. 97;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        110 TKPPR 114


RESULT 4
Q9PGQ8
ID   Q9PGQ8      PRELIMINARY;      PRT;   142 AA.
AC   Q9PGQ8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein Xf0240.
GN   XF0240.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9a5c;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
DR   EMBL; AE003878; AAF83053.1; -.
DR   PIR; C82829; C82829.
DR   InterPro; IPR002577; DUF24.
DR   Pfam; PF01638; DUF24; 1.
DR   ProDom; PD004032; DUF24; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   142 AA;  15878 MW;  CAD3A2B6EDB13E31 CRC64;

  Query Match             100.0%;  Score 29;  DB 16;  Length 142;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         87 TKPPR 91


RESULT 5
Q87EU9
ID   Q87EU9      PRELIMINARY;      PRT;   142 AA.
AC   Q87EU9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Conserved hypothetical protein.
GN   PD0199.
OS   Xylella fastidiosa (strain Temecula1 / ATCC 700964).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xylella.
OX   NCBI_TaxID=183190;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22421331; PubMed=12533478;
RA   Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA   Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA   Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA   Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorry H., Tsai S.M.,
RA   Carrer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA   Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA   Marino C.L., Giglioti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA   Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA   da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA   Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sassaki F.T., Sena J.A.D.,
RA   de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA   Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA   Kitajima J.P.;
RT   "Comparative analyses of the complete genome sequences of Pierce's
RT   disease and citrus variegated chlorosis strains of Xylella
RT   fastidiosa.";
RL   J. Bacteriol. 185:1018-1026(2003).
DR   EMBL; AE012553; AAO28090.1; -.
DR   InterPro; IPR002577; DUF24.
DR   Pfam; PF01638; DUF24; 1.
```

```
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   142 AA;  15808 MW;  D52ED9B736C13E31 CRC64;

  Query Match           100.0%;  Score 29;  DB 16;  Length 142;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         87 TKPPR 91


RESULT 6
Q9I996
ID   Q9I996      PRELIMINARY;      PRT;   143 AA.
AC   Q9I996;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Thyroid-stimulating hormone precursor.
GN   TSH.
OS   Acipenser baerii (Siberian sturgeon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC   Acipenser.
OX   NCBI_TaxID=27689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RX   MEDLINE=20318422; PubMed=10859263;
RA   Querat B., Sellouk A., Salmon C.;
RT   "Phylogenetic analysis of the vertebrate glycoprotein hormone family
RT   including new sequences of sturgeon (Acipenser baeri) subunits of the
RT   two gonadotropins and the thyroid stimulating hormone.";
RL   Biol. Reprod. 63:222-228(2000).
DR   EMBL; AJ251659; CAB93505.1; -.
DR   HSSP; P01233; 1HCN.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005179; F:hormone activity; IEA.
DR   InterPro; IPR006208; Cys_knot.
DR   InterPro; IPR001545; Gly_hormoneB.
DR   Pfam; PF00007; Cys_knot; 1.
DR   SMART; SM00068; GHB; 1.
DR   PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR   PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
KW   Signal.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    143       THYROID-STIMULATING HORMONE.
SQ   SEQUENCE   143 AA;  15566 MW;  B79E009F7F0ED315 CRC64;

  Query Match           100.0%;  Score 29;  DB 13;  Length 143;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        126 TKPPR 130


RESULT 7
Q7T922
ID   Q7T922      PRELIMINARY;      PRT;   169 AA.
AC   Q7T922;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Agnoprotein.
OS   Human adenovirus B.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=108098;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=35p;
RX   PubMed=12857895;
RA   Vogels R., Zuijdgeest D., Van Rijnsoever R., Hartkoorn E., Damen I.,
RA   De Bethune M.P., Kostense S., Penders G., Helmus N., Koudstaal W.,
RA   Cecchini M., Wetterwald A., Sprangers M., Lemckert A., Ophorst O.,
RA   Koel B., Van Meerendonk M., Quax P., Panitti L., Grimbergen J.,
RA   Bout A., Goudsmit J., Havenga M.;
RT   "Replication-Deficient Human Adenovirus Type 35 Vectors for Gene
RT   Transfer and Vaccination: Efficient Human Cell Infection and Bypass of
RT   Preexisting Adenovirus Immunity.";
RL   J. Virol. 77:8263-8271(2003).
DR   EMBL; AY271307; AAP92368.1; -.
SQ   SEQUENCE   169 AA;  18824 MW;  DC18351F66E90119 CRC64;

  Query Match           100.0%;  Score 29;  DB 12;  Length 169;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         97 TKPPR 101


RESULT 8
Q8F4V4
ID   Q8F4V4      PRELIMINARY;      PRT;   190 AA.
AC   Q8F4V4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Conserved hypothetical protein with UVR motif.
GN   LA1935.
OS   Leptospira interrogans.
OC   Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX   NCBI_TaxID=173;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA   Ren S.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE011366; AAN49134.1; -.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0004518; F:nuclease activity; IEA.
DR   GO; GO:0006289; P:nucleotide-excision repair; IEA.
DR   InterPro; IPR003729; DUF151.
DR   InterPro; IPR001943; UvrB/C.
DR   Pfam; PF02577; DUF151; 1.
DR   Pfam; PF02151; UVR; 1.
DR   PROSITE; PS50151; UVR; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   190 AA;  21142 MW;  9A13EB59D646E0AD CRC64;

  Query Match           100.0%;  Score 29;  DB 16;  Length 190;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         52 TKPPR 56


RESULT 9
Q9M250
ID   Q9M250      PRELIMINARY;      PRT;   191 AA.
AC   Q9M250;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   F7M19_120.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
```

```
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nyakatura G., Fartmann B., Dauner D., Sterr W., Holland R.,
RA   Weichselgartner M., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.,
RA   Quetier F., Salanoubat M.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL138643; CAB86482.1; -.
DR   PIR; T47369; T47369.
KW   Hypothetical protein.
SQ   SEQUENCE   191 AA;  21741 MW;  26F8764BEECBA85D CRC64;

  Query Match          100.0%;  Score 29;  DB 10;  Length 191;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         78 TKPPR 82


RESULT 10
Q98674
ID   Q98674      PRELIMINARY;      PRT;   198 AA.
AC   Q98674;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   UORF3.
OS   Simian cytomegalovirus.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Betaherpesvirinae; Cytomegalovirus.
OX   NCBI_TaxID=10364;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=11725047;
RA   Chang Y., Jeang K., Lietman T., Hayward G.S.;
RT   "Structural Organization of the Spliced Immediate-Early Gene Complex
RT   that Encodes the Major Acidic Nuclear (IE1) and Transactivator (IE2)
RT   Proteins of African Green Monkey Cytomegalovirus.";
RL   J. Biomed. Sci. 2:105-130(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90080130; PubMed=2152815;
RA   Chang Y.N., Crawford S., Stall J., Rawlins D.R., Jeang K.T.,
RA   Hayward G.S.;
RT   "The palindromic series I repeats in the simian cytomegalovirus major
RT   immediate-early promoter behave as both strong basal enhancers and
RT   cyclic AMP response elements.";
RL   J. Virol. 64:264-277(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87198884; PubMed=3033283;
RA   Jeang K.T., Rawlins D.R., Rosenfeld P.J., Shero J.H., Kelly T.J.,
RA   Hayward G.S.;
RT   "Multiple tandemly repeated binding sites for cellular nuclear factor
RT   1 that surround the major immediate-early promoters of simian and
RT   human cytomegalovirus.";
RL   J. Virol. 61:1559-1570(1987).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93100836; PubMed=8380090;
RA   Chang Y.N., Jeang K.T., Chiou C.J., Chan Y.J., Pizzorno M.,
RA   Hayward G.S.;
RT   "Identification of a large bent DNA domain and binding sites for serum
RT   response factor adjacent to the NFI repeat cluster and enhancer region
RT   in the major IE94 promoter from simian cytomegalovirus.";
RL   J. Virol. 67:516-529(1993).
DR   EMBL; U38308; AAB16873.1; -.
SQ   SEQUENCE   198 AA;  22221 MW;  DB6CE552D3775B0A CRC64;

  Query Match          100.0%;  Score 29;  DB 12;  Length 198;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        193 TKPPR 197


RESULT 11
Q8H435
ID   Q8H435      PRELIMINARY;      PRT;   212 AA.
AC   Q8H435;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   P0407H12.35 protein.
GN   P0407H12.35.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
RT   clone:P0407H12.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP004303; BAC21456.1; -.
SQ   SEQUENCE   212 AA;  24492 MW;  0EB7FD1860C63994 CRC64;

  Query Match          100.0%;  Score 29;  DB 10;  Length 212;
  Best Local Similarity  100.0%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        108 TKPPR 112


RESULT 12
Q8U6C3
ID   Q8U6C3      PRELIMINARY;      PRT;   222 AA.
AC   Q8U6C3;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein Atu4884.
GN   ATU4884 OR AGR_L_19.
OS   Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX   NCBI_TaxID=176299;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608550; PubMed=11743193;
RA   Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA   Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA   Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA   Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA   Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA   Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA   Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA   Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA   Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA   Nester E.W.;
RT   "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT   C58.";
RL   Science 294:2317-2323(2001).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608551; PubMed=11743194;
RA   Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA   Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA   Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA   Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA   Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA   Cielo C., Slater S.;
RT   "Genome sequence of the plant pathogen and biotechnology agent
RT   Agrobacterium tumefaciens C58.";
RL   Science 294:2323-2328(2001).
DR   EMBL; AE009416; AAL45678.1; ALT_INIT.
DR   EMBL; AE008198; AAK88561.1; -.
DR   PIR; AH3157; AH3157.
DR   PIR; G98129; G98129.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR002145; HTH_CopG.
DR   Pfam; PF01402; HTH_4; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   222 AA;  24796 MW;  C8E8A5C3043B6EF3 CRC64;

  Query Match           100.0%;  Score 29;  DB 16;  Length 222;
  Best Local Similarity  100.0%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         77 TKPPR 81


RESULT 13
O45855
ID   O45855      PRELIMINARY;      PRT;   226 AA.
AC   O45855;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   T27E7.1 protein.
GN   T27E7.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cummings P.N.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z82284; CAB05287.1; -.
DR   PIR; T25360; T25360.
DR   WormPep; T27E7.1; CE16505.
SQ   SEQUENCE   226 AA;  25159 MW;  1C98B5A2873B6737 CRC64;

  Query Match           100.0%;  Score 29;  DB 5;  Length 226;
  Best Local Similarity  100.0%;  Pred. No. 1.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         75 TKPPR 79


RESULT 14
Q810F4
ID   Q810F4      PRELIMINARY;      PRT;   227 AA.
AC   Q810F4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   FAM3C-like protein.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Bone marrow;
RA   Buki K.G., Vaananen K.;
RT   "Novel genes in rat bone marrow.";
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY228475; AAO73558.1; -.
SQ   SEQUENCE   227 AA;  24713 MW;  43A84E5B3266D44A CRC64;

  Query Match           100.0%;  Score 29;  DB 11;  Length 227;
  Best Local Similarity  100.0%;  Pred. No. 1.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         51 TKPPR 55


RESULT 15
Q7TML1
ID   Q7TML1      PRELIMINARY;      PRT;   227 AA.
AC   Q7TML1;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   D6Wsu176e protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Breast tumor;
RA   Strausberg R.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC055853; AAH55853.1; -.
SQ   SEQUENCE   227 AA;  24783 MW;  98F02167EBE68CE1 CRC64;

  Query Match           100.0%;  Score 29;  DB 11;  Length 227;
  Best Local Similarity  100.0%;  Pred. No. 1.9e+02;
```

```
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         51 TKPPR 55


RESULT 16
Q8SA84
ID   Q8SA84      PRELIMINARY;      PRT;   256 AA.
AC   Q8SA84;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Z195D10.20 protein (Fragment).
GN   Z195D10.20.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. B73;
RA   Ramakrishna W., Emberton J., SanMiguel P., Bennetzen J.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. B73;
RA   Doebley J.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. B73;
RA   Llaca V., Linton E.W., Young S., Kovchok S., Messing J.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF466646; AAL76008.1; -.
DR   InterPro; IPR005333; TCP.
DR   Pfam; PF03634; TCP; 1.
FT   NON_TER     256    256
SQ   SEQUENCE   256 AA;  27971 MW;  79596CC55483648E CRC64;

  Query Match             100.0%;  Score 29;  DB 10;  Length 256;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        216 TKPPR 220


RESULT 17
O35455
ID   O35455      PRELIMINARY;      PRT;   260 AA.
AC   O35455;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Homeobox protein Nkx2.6 (Fragment).
GN   NKX2-6 OR NKX2.6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98145949; PubMed=9486544;
RA   Nikolova M., Chen X., Lufkin T.;
RT   "Nkx2.6 expression is transiently and specifically restricted to the
RT   branchial region of pharyngeal-stage mouse embryos.";
RL   Mech. Dev. 69:215-218(1997).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; AF030113; AAB86406.1; -.
DR   HSSP; P23441; 1FTT.
DR   MGD; MGI:97351; Nkx2-6.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1      1
SQ   SEQUENCE   260 AA;  28360 MW;  3264612FA98EBF9B CRC64;

  Query Match             100.0%;  Score 29;  DB 11;  Length 260;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        161 TKPPR 165


RESULT 18
P74248
ID   P74248      PRELIMINARY;      PRT;   278 AA.
AC   P74248;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein slr1169.
GN   SLR1169.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
DR   EMBL; D90913; BAA18342.1; -.
DR   PIR; S75883; S75883.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   278 AA;  30966 MW;  E6503E98DB8C81EA CRC64;

  Query Match             100.0%;  Score 29;  DB 16;  Length 278;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        240 TKPPR 244


RESULT 19
Q24859
ID   Q24859      PRELIMINARY;      PRT;   283 AA.
AC   Q24859;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Entamoeba histolytica.
OC   Eukaryota; Entamoebidae; Entamoeba.
```

```
OX   NCBI_TaxID=5759;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HM-1:IMSS;
RX   MEDLINE=95327678; PubMed=7604025;
RA   Clark C.G., Roger A.J.;
RT   "Direct evidence for secondary loss of mitochondria in Entamoeba
RT   histolytica.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:6518-6521(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HM-1:IMSS;
RA   Clark C.G.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L39933; AAC41578.1; -.
DR   PIR; T18299; T18299.
KW   Hypothetical protein.
FT   NON_TER     283     283
SQ   SEQUENCE   283 AA;  32222 MW;  DD43E869752F9697 CRC64;

  Query Match             100.0%;  Score 29;  DB 5;  Length 283;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        208 TKPPR 212


RESULT 20
O60440
ID   O60440      PRELIMINARY;      PRT;   316 AA.
AC   O60440;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Alpha-actinin-2 associated LIM protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Xia H., Bredt D.S.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 LIM DOMAIN. THE LIM DOMAIN BINDS 2 ZINC
CC       IONS.
CC   -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
DR   EMBL; AF002282; AAC16672.1; -.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   InterPro; IPR001781; LIM.
DR   InterPro; IPR001478; PDZ.
DR   InterPro; IPR006643; ZASP.
DR   Pfam; PF00412; LIM; 1.
DR   Pfam; PF00595; PDZ; 1.
DR   ProDom; PD000094; LIM; 1.
DR   SMART; SM00132; LIM; 1.
DR   SMART; SM00228; PDZ; 1.
DR   SMART; SM00735; ZM; 1.
DR   PROSITE; PS50023; LIM_DOMAIN_2; 1.
DR   PROSITE; PS50106; PDZ; 1.
KW   LIM domain; Metal-binding; Zinc.
SQ   SEQUENCE   316 AA;  34266 MW;  C6D0A5EA5DA3FF90 CRC64;

  Query Match             100.0%;  Score 29;  DB 4;  Length 316;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        301 TKPPR 305


RESULT 21
Q9FLJ1
ID   Q9FLJ1      PRELIMINARY;      PRT;   325 AA.
AC   Q9FLJ1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Genomic DNA, chromosome 5, P1 clone:MIO24.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RX   MEDLINE=98290546; PubMed=9628582;
RA   Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
RT   Sequence features of the regions of 1,456,315 bp covered by nineteen
RT   physically assigned P1 and TAC clones.";
RL   DNA Res. 5:41-54(1998).
DR   EMBL; AB010074; BAB11237.1; -.
SQ   SEQUENCE   325 AA;  35789 MW;  9887D8FDFA81C6AA CRC64;

  Query Match             100.0%;  Score 29;  DB 10;  Length 325;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          9 TKPPR 13


RESULT 22
Q9FL70
ID   Q9FL70      PRELIMINARY;      PRT;   326 AA.
AC   Q9FL70;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cotton fiber expressed protein 1-like protein.
GN   AT5G54300.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RX   MEDLINE=98344145; PubMed=9679202;
RA   Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT   features of the regions of 1,381,565 bp covered by twenty one
RT   physically assigned P1 and TAC clones.";
RL   DNA Res. 5:131-145(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA   Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA   Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA   Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA   Ecker J., Theologis A., Davis R.W.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA   Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA   Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
```

```
RA   Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA   Ecker J., Theologis A., Davis R.W.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB010695; BAB10753.1; -.
DR   EMBL; AY099819; AAM20670.1; -.
DR   EMBL; BT000327; AAN15646.1; -.
DR   InterPro; IPR008480; DUF761.
DR   Pfam; PF05553; DUF761; 1.
SQ   SEQUENCE   326 AA;  36448 MW;  6EE00BC2D23B1F5F CRC64;

  Query Match             100.0%;  Score 29;  DB 10;  Length 326;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        210 TKPPR 214


RESULT 23
Q8CAP5
ID   Q8CAP5      PRELIMINARY;      PRT;   334 AA.
AC   Q8CAP5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical GRAM domain containing protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK038289; BAC29960.1; -.
DR   InterPro; IPR004182; GRAM_dom.
DR   Pfam; PF02893; GRAM; 1.
DR   SMART; SM00568; GRAM; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   334 AA;  37116 MW;  57AE3045256D81A9 CRC64;

  Query Match             100.0%;  Score 29;  DB 11;  Length 334;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        306 TKPPR 310


RESULT 24
Q8UKK5
ID   Q8UKK5      PRELIMINARY;      PRT;   350 AA.
AC   Q8UKK5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   ATP-dependent DNA ligase.
GN   ATU5097 OR AGR_PAT_142.
OS   Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OG   Plasmid AT.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX   NCBI_TaxID=176299;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608550; PubMed=11743193;
RA   Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA   Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA   Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA   Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA   Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA   Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA   Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA   Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA   Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA   Nester E.W.;
RT   "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT   C58.";
RL   Science 294:2317-2323(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608551; PubMed=11743194;
RA   Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA   Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA   Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA   Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA   Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA   Cielo C., Slater S.;
RT   "Genome sequence of the plant pathogen and biotechnology agent
RT   Agrobacterium tumefaciens C58.";
RL   Science 294:2323-2328(2001).
DR   EMBL; AE008934; AAL45787.1; -.
DR   EMBL; AE007882; AAK90473.1; -.
DR   PIR; AE3171; AE3171.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003910; F:DNA ligase (ATP) activity; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0006310; P:DNA recombination; IEA.
DR   GO; GO:0006281; P:DNA repair; IEA.
DR   GO; GO:0006260; P:DNA replication; IEA.
DR   InterPro; IPR000977; DNA_ligase.
DR   Pfam; PF01068; DNA_ligase; 1.
DR   Pfam; PF04679; DNA_ligase_A_C; 1.
DR   PROSITE; PS50160; DNA_LIGASE_A3; 1.
KW   Ligase; Plasmid; Complete proteome.
SQ   SEQUENCE   350 AA;  39363 MW;  32C15E7081D8233E CRC64;

  Query Match             100.0%;  Score 29;  DB 16;  Length 350;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          2 TKPPR 6


RESULT 25
Q23287
ID   Q23287      PRELIMINARY;      PRT;   351 AA.
AC   Q23287;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein (PIP-1).
GN   ZC404.8 OR PIP-1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Bentley D., Le T.T.;
RT   "The sequence of C. elegans cosmid ZC404.";
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Ogura K., Kishimoto N., Mitani S., Gengyo-Ando K., Kohara Y.;
RT   "Translational control of maternal glp-1 mRNA by POS-1 and its
RT   interacting protein SPN-4 in Caenorhabditis elegans.";
RL   Development 0:0-0(2003).
DR   EMBL; U55363; AAA97963.1; -.
DR   EMBL; AB052819; BAC65239.1; -.
DR   PIR; T29369; T29369.
DR   HSSP; P09651; 1HA1.
DR   WormPep; ZC404.8; CE07598.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00076; rrm; 1.
DR   SMART; SM00360; RRM; 1.
DR   PROSITE; PS50102; RRM; 1.
DR   PROSITE; PS00030; RRM_RNP_1; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   351 AA;  39213 MW;  A83B7F159472851E CRC64;

  Query Match             100.0%;  Score 29;  DB 5;  Length 351;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        199 TKPPR 203


RESULT 26
O60439
ID   O60439      PRELIMINARY;      PRT;   364 AA.
AC   O60439;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Alpha-actinin-2 associated LIM protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skeletal muscle;
RA   Xia H., Bredt D.S.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 LIM DOMAIN. THE LIM DOMAIN BINDS 2 ZINC
CC       IONS.
CC   -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
DR   EMBL; AF002280; AAC16670.1; -.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   InterPro; IPR001781; LIM.
DR   InterPro; IPR001478; PDZ.
DR   InterPro; IPR006643; ZASP.
DR   Pfam; PF00412; LIM; 1.
DR   Pfam; PF00595; PDZ; 1.
DR   ProDom; PD000094; LIM; 1.
DR   SMART; SM00132; LIM; 1.
DR   SMART; SM00228; PDZ; 1.
DR   SMART; SM00735; ZM; 1.
DR   PROSITE; PS50023; LIM_DOMAIN_2; 1.
DR   PROSITE; PS50106; PDZ; 1.
KW   LIM domain; Metal-binding; Zinc.
SQ   SEQUENCE   364 AA;  39231 MW;  EAAA2A75466D3E12 CRC64;

  Query Match             100.0%;  Score 29;  DB 4;  Length 364;
  Best Local Similarity   100.0%;  Pred. No. 3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        349 TKPPR 353


RESULT 27
Q8ZG89
ID   Q8ZG89      PRELIMINARY;      PRT;   370 AA.
AC   Q8ZG89;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative iron-sulfur binding protein (Hypothetical protein).
GN   YPO1417 OR Y2752.
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CO-92 / Biovar Orientalis;
RX   MEDLINE=21470413; PubMed=11586360;
RA   Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA   Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA   Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA   Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA   Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT   "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL   Nature 413:523-527(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KIM5 / Biovar Mediaevalis;
RX   MEDLINE=22137863; PubMed=12142430;
RA   Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA   Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA   Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA   Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA   Perry R.D.;
RT   "Genome sequence of Yersinia pestis KIM.";
RL   J. Bacteriol. 184:4601-4611(2002).
DR   EMBL; AJ414148; CAC90246.1; -.
DR   EMBL; AE013878; AAM86304.1; -.
DR   PIR; AC0173; AC0173.
DR   GO; GO:0005489; F:electron transporter activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR006058; 2Fe2S_fd_BS.
DR   InterPro; IPR001041; Ferredoxin.
DR   InterPro; IPR005302; MOSC.
DR   InterPro; IPR005303; MOSC_N.
DR   Pfam; PF00111; fer2; 1.
DR   Pfam; PF03473; MOSC; 1.
DR   Pfam; PF03476; MOSC_N; 1.
DR   PROSITE; PS00197; 2FE2S_FERREDOXIN; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   370 AA;  40747 MW;  CF6246C4DC713138 CRC64;

  Query Match             100.0%;  Score 29;  DB 16;  Length 370;
  Best Local Similarity   100.0%;  Pred. No. 3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        266 TKPPR 270
```

```
RESULT 28
O80865
ID   O80865      PRELIMINARY;      PRT;   375 AA.
AC   O80865;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   AT2G31010.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR   EMBL; AC004669; AAC20735.1; -.
DR   PIR; D84715; D84715.
DR   HSSP; P12931; 1FMK.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Hypothetical protein; ATP-binding; Kinase;
KW   Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   375 AA;  42588 MW;  C1F40CA3A3DE77C5 CRC64;

  Query Match             100.0%;  Score 29;  DB 10;  Length 375;
  Best Local Similarity   100.0%;  Pred. No. 3.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        121 TKPPR 125


RESULT 29
Q82NT5
ID   Q82NT5      PRELIMINARY;      PRT;   394 AA.
AC   Q82NT5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   SAV1206.
OS   Streptomyces avermitilis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=33903;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=21477403; PubMed=11572948;
RA   Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA   Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=22608306; PubMed=12692562;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industrial
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
DR   EMBL; AP005025; BAC68916.1; -.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR006090; Acyl-CoA_dh_C.
DR   Pfam; PF00441; Acyl-CoA_dh; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   394 AA;  42238 MW;  B8CF41D0C94F29D2 CRC64;

  Query Match             100.0%;  Score 29;  DB 16;  Length 394;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        386 TKPPR 390


RESULT 30
Q89JB7
ID   Q89JB7      PRELIMINARY;      PRT;   405 AA.
AC   Q89JB7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Bll5366 protein.
GN   BLL5366.
OS   Bradyrhizobium japonicum.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Bradyrhizobium.
OX   NCBI_TaxID=375;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=USDA 110;
RX   MEDLINE=22484998; PubMed=12597275;
RA   Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA   Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA   Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA   Tabata S.;
RT   "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT   Bradyrhizobium japonicum USDA110.";
RL   DNA Res. 9:189-197(2002).
DR   EMBL; AP005954; BAC50631.1; -.
DR   GO; GO:0004730; F:pseudouridylate synthase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   InterPro; IPR006145; PseudoU_synth.
DR   InterPro; IPR006224; Rlu_synth.
DR   InterPro; IPR002942; S4.
```

```
DR   Pfam; PF00849; PseudoU_synth_2; 1.
DR   Pfam; PF01479; S4; 1.
DR   PROSITE; PS01129; PSI_RLU; 1.
DR   PROSITE; PS50889; S4; 1.
KW   Complete proteome.
SQ   SEQUENCE   405 AA;  45209 MW;  2BEE412B47EE36F5 CRC64;

  Query Match           100.0%;  Score 29;  DB 16;  Length 405;
  Best Local Similarity  100.0%;  Pred. No. 3.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         42 TKPPR 46


RESULT 31
Q83A13
ID   Q83A13      PRELIMINARY;      PRT;   406 AA.
AC   Q83A13;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Site specific recombinase, phage integrase family.
GN   CBUA0010.
OS   Coxiella burnetii.
OG   Plasmid pQpH1.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC   Coxiellaceae; Coxiella.
OX   NCBI_TaxID=777;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nine Mile phase I / RSA 493;
RX   MEDLINE=22608657; PubMed=12704232;
RA   Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA   Nelson W.C., Ward N.L., Tettelin H., Davidsen T.M., Beanan M.J.,
RA   DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
RA   Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA   Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;
RT   "Complete genome sequence of the Q-fever pathogen, Coxiella
RT   burnetii.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
DR   EMBL; AE016829; AAO91587.1; -.
DR   TIGR; CBUA0010; -.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0015074; P:DNA integration; IEA.
DR   GO; GO:0006310; P:DNA recombination; IEA.
DR   InterPro; IPR002104; Phage_integrase.
DR   Pfam; PF00589; Phage_integrase; 1.
KW   Plasmid; Complete proteome.
SQ   SEQUENCE   406 AA;  47359 MW;  1D40014C4500A23E CRC64;

  Query Match           100.0%;  Score 29;  DB 16;  Length 406;
  Best Local Similarity  100.0%;  Pred. No. 3.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         98 TKPPR 102


RESULT 32
Q9D9R4
ID   Q9D9R4      PRELIMINARY;      PRT;   409 AA.
AC   Q9D9R4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   1700030J22Rik protein.
GN   1700030J22RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK006553; BAB24650.1; -.
DR   MGD; MGI:1916778; 1700030J22Rik.
SQ   SEQUENCE   409 AA;  45022 MW;  A61F0B8F9A15A802 CRC64;

  Query Match           100.0%;  Score 29;  DB 11;  Length 409;
  Best Local Similarity  100.0%;  Pred. No. 3.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        358 TKPPR 362


RESULT 33
Q8BHB7
ID   Q8BHB7      PRELIMINARY;      PRT;   409 AA.
AC   Q8BHB7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein.
GN   1700030J22RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK030075; BAC26769.1; -.
DR   EMBL; AK031529; BAC27437.1; -.
DR   MGD; MGI:1916778; 1700030J22Rik.
KW   Hypothetical protein.
SQ   SEQUENCE   409 AA;  45038 MW;  360D066F97F85A1D CRC64;

  Query Match           100.0%;  Score 29;  DB 11;  Length 409;
  Best Local Similarity  100.0%;  Pred. No. 3.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
```

```
Db        358 TKPPR 362


RESULT 34
Q45945
ID   Q45945      PRELIMINARY;      PRT;   410 AA.
AC   Q45945; Q45901;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   ORF 410.
OS   Coxiella burnetii.
OG   Plasmid QpH1.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC   Coxiellaceae; Coxiella.
OX   NCBI_TaxID=777;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NINE MILE PHASE I;
RX   MEDLINE=95145653; PubMed=7843345;
RA   Thiele D., Willems H., Haas M., Krauss H.;
RT   "Analysis of the entire nucleotide sequence of the cryptic plasmid
RT   QpH1 from Coxiella burnetti.";
RL   Eur. J. Epidemiol. 10:413-420(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NINE MILE PHASE I;
RA   Thiele D.;
RL   Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SCURRY Q217;
RA   Ritter M., Thiele D., Willems H.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X75356; CAA53126.1; -.
DR   EMBL; X93204; CAA63678.1; -.
DR   PIR; S38238; S38238.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0015074; P:DNA integration; IEA.
DR   GO; GO:0006310; P:DNA recombination; IEA.
DR   InterPro; IPR002104; Phage_integrase.
DR   Pfam; PF00589; Phage_integrase; 1.
KW   Plasmid.
SQ   SEQUENCE   410 AA;  47829 MW;  9A448C3714B330C2 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        102 TKPPR 106


RESULT 35
O52878
ID   O52878      PRELIMINARY;      PRT;   410 AA.
AC   O52878;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Coxiella burnetii.
OG   Plasmid QpRS, and Plasmid QpDV.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC   Coxiellaceae; Coxiella.
OX   NCBI_TaxID=777;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PRISCILLA Q177;
RA   Lautenschlaeger S., Jaeger C., Willems H., Baljer G.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1140;
RA   Radomski K.U., Willems H., Lautenschlaeger S., Jaeger C., Baljer G.;
RT   "Sequence of QpDV plasmid.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y15898; CAA75839.1; -.
DR   EMBL; AF131076; AAD33493.1; -.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0015074; P:DNA integration; IEA.
DR   GO; GO:0006310; P:DNA recombination; IEA.
DR   InterPro; IPR002104; Phage_integrase.
DR   Pfam; PF00589; Phage_integrase; 1.
KW   Hypothetical protein; Plasmid.
SQ   SEQUENCE   410 AA;  47828 MW;  7A49E15719D330CA CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        102 TKPPR 106


RESULT 36
Q9XIF9
ID   Q9XIF9      PRELIMINARY;      PRT;   415 AA.
AC   Q9XIF9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   T10P12.9 protein.
GN   T10P12.9.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA   Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA   Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA   Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA   Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC007203; AAD39275.1; -.
DR   PIR; F96499; F96499.
DR   HSSP; P50586; 1C8Z.
DR   InterPro; IPR000007; Tubby.
DR   Pfam; PF01167; Tub; 1.
DR   PRINTS; PR01573; SUPERTUBBY.
DR   PROSITE; PS01200; TUB_1; 1.
SQ   SEQUENCE   415 AA;  46492 MW;  38774871B95D8770 CRC64;

  Query Match             100.0%;  Score 29;  DB 10;  Length 415;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        309 TKPPR 313


RESULT 37
Q8S8E1
ID   Q8S8E1      PRELIMINARY;      PRT;   415 AA.
AC   Q8S8E1;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
```

```
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   AT2G40420.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA   Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA   Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA   Fraser C.M., Venter J.C.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC007020; AAM15348.1; -.
DR   PIR; C84829; C84829.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR   GO; GO:0006865; P:amino acid transport; IEA.
DR   InterPro; IPR002422; AA/rel_permease2.
DR   Pfam; PF01490; Aa_trans; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   415 AA;  44930 MW;  5A7B2E9C048F5A44 CRC64;

  Query Match             100.0%;  Score 29;  DB 10;  Length 415;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        169 TKPPR 173


RESULT 38
Q96Z07
ID   Q96Z07      PRELIMINARY;      PRT;   415 AA.
AC   Q96Z07;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative hyaluronan synthase.
GN   ST2021.
OS   Sulfolobus tokodaii.
OC   Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC   Sulfolobus.
OX   NCBI_TaxID=111955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JCM 10545 / 7;
RX   MEDLINE=21456156; PubMed=11572479;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA   Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA   Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA   Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA   Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA   Oshima T., Kikuchi H.;
RT   "Complete genome sequence of an aerobic thermoacidophilic
RT   Crenarchaeon, Sulfolobus tokodaii strain7.";
RL   DNA Res. 8:123-140(2001).
DR   EMBL; AP000988; BAB67119.1; -.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF00535; Glycos_transf_2; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   415 AA;  48922 MW;  A155BE3BF391D822 CRC64;

  Query Match             100.0%;  Score 29;  DB 17;  Length 415;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        259 TKPPR 263


RESULT 39
Q944S3
ID   Q944S3      PRELIMINARY;      PRT;   429 AA.
AC   Q944S3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   At1g43640/T10P12_16 (Hypothetical protein).
GN   AT1G43640.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA   Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA   Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA   Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA   Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA   Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA   Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA   Ecker J.R.;
RT   "Arabidopsis cDNA clones.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA   Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA   Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA   Davis R.W., Ecker J.R., Theologis A.;
RT   "Full Length cDNA of gene At1g43640 (GI:15218431).";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W.,
RA   Lee J.M., Quach H.L., Tang C., Toriumi M., Wu H.C., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA   Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA   Theologis A.;
RT   "Arabidopsis Open Reading Frame (ORF) Clones.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF424565; AAL11559.1; -.
DR   EMBL; AY072154; AAL59976.1; -.
DR   EMBL; AY122972; AAM67505.1; -.
DR   InterPro; IPR001810; F-box.
DR   InterPro; IPR000007; Tubby.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF01167; Tub; 1.
DR   PRINTS; PR01573; SUPERTUBBY.
DR   PROSITE; PS01200; TUB_1; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   429 AA;  47845 MW;  DC1E315A4C202556 CRC64;

  Query Match             100.0%;  Score 29;  DB 10;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        323 TKPPR 327
```

```
RESULT 40
Q9D2W5
ID   Q9D2W5      PRELIMINARY;      PRT;   432 AA.
AC   Q9D2W5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   9130427A09Rik protein.
GN   9130427A09RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cecum;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK018691; BAB31347.1; -.
DR   MGD; MGI:1914815; 9130427A09Rik.
DR   InterPro; IPR004182; GRAM_dom.
DR   Pfam; PF02893; GRAM; 1.
DR   SMART; SM00568; GRAM; 1.
SQ   SEQUENCE   432 AA;  47932 MW;  2D9898518F2B544B CRC64;

  Query Match           100.0%;  Score 29;  DB 11;  Length 432;
  Best Local Similarity  100.0%;  Pred. No. 3.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        300 TKPPR 304


RESULT 41
Q06349
ID   Q06349      PRELIMINARY;      PRT;   442 AA.
AC   Q06349;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chromosome IV COSMID 9481.
GN   YDR370C OR D9481.14.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RA   Ding H.;
RL   Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RA   Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA   Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
RA   Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
RA   Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,
RA   Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,
RA   Taich A., Trevaskis E., Vignati D., Wilcox L., Wohldman P., Vaudin M.,
RA   Wilson R., Waterston R.;
RL   Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RA   Waterston R.;
RL   Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RA   Jia Y., Cherry J.M.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U28373; AAB64806.1; -.
DR   PIR; S61165; S61165.
DR   SGD; S0002778; YDR370C.
SQ   SEQUENCE   442 AA;  50492 MW;  504F7C6AA40B50E4 CRC64;

  Query Match           100.0%;  Score 29;  DB 3;  Length 442;
  Best Local Similarity  100.0%;  Pred. No. 3.6e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         81 TKPPR 85


RESULT 42
Q9BLQ2
ID   Q9BLQ2      PRELIMINARY;      PRT;   458 AA.
AC   Q9BLQ2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   Putative avirulence protein precursor.
GN   MAP-1.
OS   Meloidogyne incognita (southern root-knot nematode).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC   Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
OX   NCBI_TaxID=6306;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21034633; PubMed=11194874;
RA   Semblat J.P., Rosso M.N., Hussey R.S., Abad P., Castagnone-Sereno P.;
RT   "Molecular cloning of a cDNA encoding an amphid-secreted putative
RT   avirulence protein from the root-knot nematode Meloidogyne
RT   incognita.";
RL   Mol. Plant Microbe Interact. 14:72-79(2001).
DR   EMBL; AJ278663; CAC27774.1; -.
DR   InterPro; IPR002965; P_rich_extensn.
DR   PRINTS; PR01217; PRICHEXTENSN.
KW   Signal.
FT   SIGNAL       1     17       POTENTIAL.
FT   CHAIN       18    458       PUTATIVE AVIRULENCE PROTEIN.
SQ   SEQUENCE   458 AA;  48951 MW;  982AD87D0818815F CRC64;

  Query Match           100.0%;  Score 29;  DB 5;  Length 458;
  Best Local Similarity  100.0%;  Pred. No. 3.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        168 TKPPR 172


RESULT 43
Q94LM4
```

```
ID   Q94LM4      PRELIMINARY;      PRT;   460 AA.
AC   Q94LM4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Unknown protein.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wing R.A., Frisch D., Presting G., Wood T., Yu Y., Soderlund C.,
RA   Kim H., Rambo T., Henry D., Simmons J.;
RT   "Rice Genomic Sequence.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC079038; AAK55772.1; -.
DR   Gramene; Q94LM4; -.
DR   InterPro; IPR008938; ARM.
SQ   SEQUENCE   460 AA;  50005 MW;  8CFCEC9231977779 CRC64;

  Query Match             100.0%;  Score 29;  DB 10;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        136 TKPPR 140


RESULT 44
Q828Y7
ID   Q828Y7      PRELIMINARY;      PRT;   491 AA.
AC   Q828Y7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative GTP-binding protein.
GN   SAV6524.
OS   Streptomyces avermitilis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=33903;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=21477403; PubMed=11572948;
RA   Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA   Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=22608306; PubMed=12692562;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industrial
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
DR   EMBL; AP005047; BAC74235.1; -.
DR   GO; GO:0000166; F:nucleotide binding; IEA.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR005289; GTP-bindding_dom.
DR   InterPro; IPR006073; GTP1_OBG.
DR   InterPro; IPR002917; MMR_HSR1.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF01926; MMR_HSR1; 1.
DR   PRINTS; PR00326; GTP1OBG.
DR   SMART; SM00382; AAA; 2.
DR   TIGRFAMs; TIGR00650; MG442; 2.
DR   TIGRFAMs; TIGR00231; small_GTP; 2.
KW   Complete proteome.
SQ   SEQUENCE   491 AA;  53353 MW;  7BADD8E0449D14AC CRC64;

  Query Match             100.0%;  Score 29;  DB 16;  Length 491;
  Best Local Similarity   100.0%;  Pred. No. 4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        440 TKPPR 444


RESULT 45
Q82GE0
ID   Q82GE0      PRELIMINARY;      PRT;   499 AA.
AC   Q82GE0;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative transport integral membrane protein.
GN   SAV3957.
OS   Streptomyces avermitilis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=33903;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=21477403; PubMed=11572948;
RA   Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA   Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=22608306; PubMed=12692562;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industrial
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
DR   EMBL; AP005037; BAC71669.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR004638; Efflux_EmrB.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR00711; efflux_EmrB; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW   Complete proteome.
SQ   SEQUENCE   499 AA;  52225 MW;  F580C79D5469F2CF CRC64;

  Query Match             100.0%;  Score 29;  DB 16;  Length 499;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        482 TKPPR 486


RESULT 46
```

```
Q7UT23
ID   Q7UT23      PRELIMINARY;      PRT;   505 AA.
AC   Q7UT23;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   4-alpha-glucanotransferase (EC 2.4.1.25).
GN   MALQ OR RB4161.
OS   Rhodopirellula baltica.
OC   Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC   Planctomycetaceae; Pirellula.
OX   NCBI_TaxID=117;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1;
RX   MEDLINE=22735913; PubMed=12835416;
RA   Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA   Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA   Schlesner H., Amann R., Reinhardt R.;
RT   "Complete genome sequence of the marine planctomycete Pirellula sp.
RT   strain 1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR   EMBL; BX294140; CAD73618.1; -.
KW   Glycosyltransferase; Transferase; Complete proteome.
SQ   SEQUENCE   505 AA;  57353 MW;  5470834CCAF15FFF CRC64;

  Query Match             100.0%;  Score 29;  DB 16;  Length 505;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        121 TKPPR 125


RESULT 47
Q9D318
ID   Q9D318      PRELIMINARY;      PRT;   514 AA.
AC   Q9D318;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   9030613F08Rik protein.
GN   9030613F08RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Colon;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK018551; BAB31270.1; -.
DR   MGD; MGI:1921790; 9030613F08Rik.
DR   InterPro; IPR004182; GRAM_dom.
DR   Pfam; PF02893; GRAM; 1.
DR   SMART; SM00568; GRAM; 1.
SQ   SEQUENCE   514 AA;  56478 MW;  C5A5042F9D228B84 CRC64;

  Query Match             100.0%;  Score 29;  DB 11;  Length 514;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        382 TKPPR 386


RESULT 48
Q9EQY1
ID   Q9EQY1      PRELIMINARY;      PRT;   524 AA.
AC   Q9EQY1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Transcription factor Elf-1.
GN   ELF-1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21077473; PubMed=11210123;
RA   Nishiyama C., Takahashi K., Nishiyama M., Okumura K., Ra C.,
RA   Ohtake Y., Yokota T.;
RT   "Polymorphism of transcription factor Elf-1 affecting its regulatory
RT   function in transcription.";
RL   Biosci. Biotechnol. Biochem. 64:2601-2607(2000).
DR   EMBL; AB030217; BAB20035.1; -.
DR   HSSP; P28324; 1BC8.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR000418; Ets.
DR   InterPro; IPR002341; HSF_ETS.
DR   Pfam; PF00178; Ets; 1.
DR   PRINTS; PR00454; ETSDOMAIN.
DR   SMART; SM00413; ETS; 1.
DR   PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR   PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR   PROSITE; PS50061; ETS_DOMAIN_3; 1.
SQ   SEQUENCE   524 AA;  56580 MW;  87834210FC7AAD1A CRC64;

  Query Match             100.0%;  Score 29;  DB 11;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 4.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db         88 TKPPR 92


RESULT 49
Q9VQ69
ID   Q9VQ69      PRELIMINARY;      PRT;   547 AA.
AC   Q9VQ69; Q961C7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CG31672 protein (LD22825p).
GN   BEST:LD15963 OR CG4248 OR CG31672.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
```

```
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003584; AAF51310.2; -.
DR   EMBL; AY051677; AAK93101.1; -.
DR   FlyBase; FBgn0028952; BEST:LD15963.
SQ   SEQUENCE   547 AA;  63439 MW;  74CFFD0A1BA4F195 CRC64;

  Query Match           100.0%;  Score 29;  DB 5;  Length 547;
  Best Local Similarity  100.0%;  Pred. No. 4.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        205 TKPPR 209


RESULT 50
Q96HX2
ID   Q96HX2      PRELIMINARY;      PRT;   574 AA.
AC   Q96HX2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye;
RA   Strausberg R.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC007998; AAH07998.1; -.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   574 AA;  60606 MW;  E598819330ADFEE4 CRC64;

  Query Match           100.0%;  Score 29;  DB 4;  Length 574;
  Best Local Similarity  100.0%;  Pred. No. 4.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db        387 TKPPR 391


Search completed: March  3, 2004, 12:18:34
Job time : 44 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 3, 2004, 12:11:12 ; Search time 55 Seconds
(without alignments)
25.686 Million cell updates/sec

Title: US-09-871-974-2
Perfect score: 29
Sequence: 1 TKPPR 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	5	2	AAW11052	Aaw11052 Leukocyte
2	29	100.0	5	2	AAW31146	Aaw31146 Amyloid p
3	29	100.0	5	2	AAY49845	Aay49845 Tuftsin r
4	29	100.0	5	5	ABB08442	Abb08442 Tuftsin r
5	29	100.0	5	5	AAM51906	Aam51906 Tertiary
6	29	100.0	5	7	ADD10684	Add10684 Tuftsin a
7	29	100.0	6	2	AAR85540	Aar85540 Tuftsin a
8	29	100.0	6	2	AAW11053	Aaw11053 Leukocyte
9	29	100.0	6	2	AAW31148	Aaw31148 Amyloid p
10	29	100.0	6	2	AAW31147	Aaw31147 Amyloid p
11	29	100.0	6	2	AAY49840	Aay49840 Tuftsin r
12	29	100.0	6	5	ABB08444	Abb08444 Tuftsin r
13	29	100.0	6	5	AAM51904	Aam51904 Tertiary
14	29	100.0	6	5	AAM51905	Aam51905 Tertiary
15	29	100.0	6	7	ADD10689	Add10689 Tuftsin a
16	29	100.0	6	7	ADD10687	Add10687 Tuftsin a
17	29	100.0	7	2	AAY49842	Aay49842 Tuftsin r
18	29	100.0	7	2	AAY49841	Aay49841 Tuftsin r
19	29	100.0	7	5	ABB08447	Abb08447 Tuftsin r
20	29	100.0	7	5	ABB08445	Abb08445 Linear an
21	29	100.0	7	7	ADD10688	Add10688 Cyclic Tu
22	29	100.0	7	7	ADD10686	Add10686 Tuftsin a
23	29	100.0	8	2	AAR88740	Aar88740 Tuftsin a
24	29	100.0	8	2	AAR76218	Aar76218 peptide-l
25	29	100.0	8	2	AAW11058	Aaw11058 Leukocyte
26	29	100.0	8	2	AAY49844	Aay49844 Tuftsin r
27	29	100.0	8	5	ABB08448	Abb08448 Tuftsin r
28	29	100.0	9	2	AAR85539	Aar85539 Metal che
29	29	100.0	9	2	AAR85535	Aar85535 Metal che
30	29	100.0	9	2	AAR88735	Aar88735 Tuftsin a
31	29	100.0	9	2	AAR88738	Aar88738 Tuftsin a
32	29	100.0	9	2	AAR88741	Aar88741 Tuftsin a
33	29	100.0	9	2	AAR76219	Aar76219 Peptide-l
34	29	100.0	9	2	AAW11055	Aaw11055 Leukocyte
35	29	100.0	9	2	AAW11059	Aaw11059 Leukocyte
36	29	100.0	9	2	AAW11054	Aaw11054 Leukocyte
37	29	100.0	9	2	AAW11056	Aaw11056 Leukocyte
38	29	100.0	9	2	AAW11057	Aaw11057 Leukocyte
39	29	100.0	9	2	AAW03420	Aaw03420 Peptide u
40	29	100.0	9	2	AAY23752	Aay23752 Peptide R
41	29	100.0	10	2	AAR88739	Aar88739 Tuftsin a
42	29	100.0	10	7	ADE81165	Ade81165 Tufstin m
43	29	100.0	10	7	ADE81166	Ade81166 Tufstin m
44	29	100.0	11	2	AAR88736	Aar88736 Tuftsin a
45	29	100.0	11	7	ADD10691	Add10691 Tuftsin a
46	29	100.0	11	7	ADD10690	Add10690 Tuftsin a
47	29	100.0	12	5	ABB08446	Abb08446 Tuftsin r
48	29	100.0	20	5	ABB08449	Abb08449 99mTc rad
49	29	100.0	53	5	ABP34343	Abp34343 Human ORF
50	29	100.0	65	4	ABG07007	Abg07007 Novel hum
51	29	100.0	99	4	AAU51890	Aau51890 Propionib
52	29	100.0	99	5	ABB84958	Abb84958 Human PRO
53	29	100.0	99	5	ABG34045	Abg34045 Human Pro
54	29	100.0	99	5	ABB95564	Abb95564 Human ang
55	29	100.0	99	6	ABM48409	Abm48409 Propionib
56	29	100.0	99	6	ADA01300	Ada01300 Human PRO
57	29	100.0	99	6	ADA43729	Ada43729 Human sec
58	29	100.0	99	6	ADA43497	Ada43497 Human sec
59	29	100.0	99	6	ADA01172	Ada01172 Human PRO
60	29	100.0	99	7	ADA01056	Ada01056 Human sec
61	29	100.0	99	7	ADA43613	Ada43613 Human sec
62	29	100.0	99	7	ADA06875	Ada06875 Human PRO
63	29	100.0	99	7	ADA08363	Ada08363 Novel hum
64	29	100.0	99	7	ADB99656	Adb99656 Human PRO
65	29	100.0	99	7	ADB86939	Adb86939 Human PRO
66	29	100.0	99	7	ADB66094	Adb66094 Human sec
67	29	100.0	99	7	ADB99772	Adb99772 Human PRO
68	29	100.0	99	7	ADB99427	Adb99427 Novel hum
69	29	100.0	99	7	ADB65978	Adb65978 Human sec
70	29	100.0	99	7	ADC23376	Adc23376 Human tra
71	29	100.0	99	7	ADC26069	Adc26069 Human PRO
72	29	100.0	99	7	ADD10573	Add10573 Human sec
73	29	100.0	99	7	ADD11533	Add11533 Human sec
74	29	100.0	99	7	ADD37326	Add37326 Human sec
75	29	100.0	99	7	ADE04896	Ade04896 Human PRO
76	29	100.0	99	7	ADE11202	Ade11202 Human PRO
77	29	100.0	99	7	ADD88133	Add88133 Human PRO
78	29	100.0	99	7	ADD95428	Add95428 Human sec
79	29	100.0	99	7	ADE06358	Ade06358 Human PRO
80	29	100.0	99	7	ADE38133	Ade38133 Human PRO
81	29	100.0	99	7	ADD88249	Add88249 Human PRO
82	29	100.0	99	7	ADD90830	Add90830 Human sec
83	29	100.0	99	8	ADE51683	Ade51683 Human sec
84	29	100.0	99	8	ADE51799	Ade51799 Human sec
85	29	100.0	99	8	ADE37657	Ade37657 Human sec
86	29	100.0	99	8	ADE37541	Ade37541 Human sec
87	29	100.0	99	8	ADD95312	Add95312 Human sec
88	29	100.0	99	8	ADE38012	Ade38012 Human PRO
89	29	100.0	99	8	ADE76101	Ade76101 Human PRO
90	29	100.0	99	8	ADE39424	Ade39424 Human PRO
91	29	100.0	99	8	ADE04228	Ade04228 Human PRO
92	29	100.0	99	8	ADE39825	Ade39825 Human PRO
93	29	100.0	99	8	ADE19690	Ade19690 Human PRO
94	29	100.0	99	8	ADE77268	Ade77268 Human sec
95	29	100.0	99	8	ADE65376	Ade65376 Human PRO
96	29	100.0	99	8	ADE75985	Ade75985 Human PRO
97	29	100.0	99	8	ADE37896	Ade37896 Human PRO
98	29	100.0	99	8	ADE64506	Ade64506 Human PRO

99	29	100.0	99	8	ADE41534	Ade41534 Human sec
100	29	100.0	99	8	ADE38841	Ade38841 Human PRO
101	29	100.0	99	8	ADE51915	Ade51915 Human sec
102	29	100.0	99	8	ADD90946	Add90946 Human sec
103	29	100.0	99	8	ADE38725	Ade38725 Human PRO
104	29	100.0	99	8	ADE37425	Ade37425 Human sec
105	29	100.0	99	8	ADE06242	Ade06242 Human PRO
106	29	100.0	99	8	ADD90101	Add90101 Human sec
107	29	100.0	99	8	ADE38609	Ade38609 Human PRO
108	29	100.0	99	8	ADE39540	Ade39540 Human PRO
109	29	100.0	99	8	ADD89145	Add89145 Human PRO
110	29	100.0	99	8	ADD88912	Add88912 Human PRO
111	29	100.0	99	8	ADE19806	Ade19806 Human PRO
112	29	100.0	99	8	ADE77384	Ade77384 Human sec
113	29	100.0	99	8	ADE65260	Ade65260 Human PRO
114	29	100.0	99	8	ADE39308	Ade39308 Human PRO
115	29	100.0	99	8	ADE38493	Ade38493 Human sec
116	29	100.0	103	4	AAU42638	Aau42638 Propionib
117	29	100.0	103	6	ABM39157	Abm39157 Propionib
118	29	100.0	107	3	AAY70535	Aay70535 Maize pla
119	29	100.0	136	4	ABG22236	Abg22236 Novel hum
120	29	100.0	149	4	ABG16691	Abg16691 Novel hum
121	29	100.0	198	4	AAU46263	Aau46263 Propionib
122	29	100.0	198	6	ABM42782	Abm42782 Propionib
123	29	100.0	216	4	ABG03775	Abg03775 Novel hum
124	29	100.0	227	5	ABG91397	Abg91397 Primate L
125	29	100.0	253	4	ABG07008	Abg07008 Novel hum
126	29	100.0	304	4	ABG16575	Abg16575 Novel hum
127	29	100.0	309	5	ABG77237	Abg77237 Selected
128	29	100.0	309	5	ABJ11108	Abj11108 Yeast sel
129	29	100.0	310	4	ABG06185	Abg06185 Novel hum
130	29	100.0	334	4	ABG17530	Abg17530 Novel hum
131	29	100.0	340	3	AAG43777	Aag43777 Arabidops
132	29	100.0	343	3	AAG43776	Aag43776 Arabidops
133	29	100.0	344	3	AAG43775	Aag43775 Arabidops
134	29	100.0	350	4	AAB50667	Aab50667 C. elegan
135	29	100.0	351	4	AAB50666	Aab50666 C. elegan
136	29	100.0	409	6	AAE34477	Aae34477 Murine AG
137	29	100.0	415	7	ADC77485	Adc77485 Sulfolobu
138	29	100.0	415	7	ADD93932	Add93932 Sulfolobu
139	29	100.0	474	4	ABB59612	Abb59612 Drosophil
140	29	100.0	494	4	ABG03323	Abg03323 Novel hum
141	29	100.0	524	3	AAB03192	Aab03192 Rat Elf-1
142	29	100.0	530	4	ABG03809	Abg03809 Novel hum
143	29	100.0	530	4	ABG01390	Abg01390 Novel hum
144	29	100.0	563	4	ABG04846	Abg04846 Novel hum
145	29	100.0	563	4	ABG16158	Abg16158 Novel hum
146	29	100.0	576	2	AAY06477	Aay06477 Human tum
147	29	100.0	576	2	AAY49546	Aay49546 Human Nsp
148	29	100.0	576	2	AAY49545	Aay49545 Human Nsp
149	29	100.0	576	2	AAY49547	Aay49547 Human Nsp
150	29	100.0	576	2	AAY49541	Aay49541 Human PRO

ALIGNMENTS

RESULT 1
AAW11052
ID AAW11052 standard; peptide; 5 AA.
XX
AC AAW11052;
XX
DT 03-JUN-1997 (first entry)
XX
DE Leukocyte-targetting peptide used in diagnostic imaging.
XX
KW Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
KW diagnostic imaging.
XX
OS Synthetic.
XX
PN WO9603427-A1.
XX
PD 08-FEB-1996.
XX
PF 28-APR-1995; 95WO-CA000249.
XX
PR 22-JUL-1994; 94US-00279155.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Pollak A, Goodbody A;
XX
DR WPI; 1996-116994/12.
XX
PT New peptide derived radionuclide chelators and metal complexes - useful
PT for diagnostic imaging.
XX
PS Claim 12; Page 19; 30pp; English.
XX
CC AAW11052 is a peptide used for targetting agents to leukocytes. This
CC peptide can be coupled to a metal radionuclide chelator (structural
CC formula given in the specification), labelled with a diagnostically
CC useful metal isotope, to form a peptide derived radionuclide chelator.
CC When the chelator is coupled to a targeting molecule and labelled with a
CC diagnostically useful metal, it can be used to detect pathological
CC conditions by diagnostic imaging. For example, AAW11052 targets the
CC chelator to leukocytes and are useful for the rapid imaging of sites of
CC local inflammation. Radionuclides used include 99mTc, 64Cu, 67Cu, 97Ru,
CC 105Rh, 109Pd, 186Re, 188Re, 198Au, 199Au, 203Pb, 212Pb and 212Bi. The
CC coupling of a targeting agent and radionuclide using a chelating agent is
CC an alternative to the direct labelling of targeting agents in which
CC radionuclides are typically bound at the more numerous low- affinity
CC sites, forming unstable complexes. The new conjugates give better
CC scintigraphic images in rat inflammation studies than known imaging
CC agents Ga-67, 99mTc-IgG, 111In-WBC and 99mTc-Nanocoll. They image more
CC rapidly than the known agents and show superior biodistribution
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 29; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5
```

RESULT 2
AAW31146
ID AAW31146 standard; peptide; 5 AA.
XX
AC AAW31146;
XX
DT 25-MAR-2003 (revised)
DT 23-JAN-1998 (first entry)
XX
DE Amyloid plaque-targeting peptide.
XX
KW Target; delivery; radionuclide chelator; diagnosis; therapy; detection;
KW atherosclerosis; thrombosis; Alzheimer's disease.
XX
OS Synthetic.
XX
PN US5659041-A.
XX
PD 19-AUG-1997.
XX
PF 02-SEP-1994; 94US-00299636.
XX
PR 19-JUL-1993; 93US-00092911.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX

PI Pollak A, Kirby RA, Dunn-Dufault R;
XX
DR WPI; 1997-424290/39.
XX
PT New thio:acetyl-aminoacid hydrazide compounds - useful as chemical
PT chelator of radionuclides for radio:imaging of target tissues of
PT diagnostic interest.
XX
PS Disclosure; Col 29; 20pp; English.
XX
CC AAW31110-W31147 are peptides used for targeting a new hydrazino-type
CC compound to various sites of disease, e.g. atherosclerotic plaque, sites
CC of infection, platelets, thrombus or amyloid plaque. The new compound is
CC a radionuclide chelator and is used to radiolabel the targeting peptides
CC for the detection and diagnostic imaging of sites of disease, e.g.
CC amyloid plaques in Alzheimer's disease. (Updated on 25-MAR-2003 to
CC correct PF field.)
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 29; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 1 TKPPR 5

RESULT 3
AAY49845
ID AAY49845 standard; peptide; 5 AA.
XX
AC AAY49845;
XX
DT 20-JAN-2000 (first entry)
XX
DE Tuftsin receptor antagonist peptide.
XX
KW Tuftsin receptor antagonist; chelate conjugate; radiopharmaceutical;
KW diagnosis; infection; inflammation; imaging; cancer; tumour.
XX
OS Synthetic.
XX
PN WO9951628-A1.
XX
PD 14-OCT-1999.
XX
PF 29-MAR-1999; 99WO-US006824.
XX
PR 03-APR-1998; 98US-0080672P.
XX
PA (DUPO) DU PONT PHARM CO.
XX
PI Edwards DS, Rajopadhye M;
XX
DR WPI; 1999-633729/54.
XX
PT New polypeptides and radiopharmacueticals used for imaging infection,
PT inflammation and cancer.
XX
PS Disclosure; Page 3; 80pp; English.
XX
CC The present invention describes polypeptide compounds of formula (I),
CC capable of direct transformation into a radiopharmaceutical: Ch-Ln-(X1-X2
CC -X3-X4-X5)d (I), where X1-X5 = amino acids; Ln = a linking group; Ch = a
CC metal bonding unit; and d is selected from 1, 2 and 3. The
CC radiopharmaceuticals are useful for the diagnosis of infection,
CC inflammation and cancer. The radiopharmaceuticals bind in vivo to the
CC tuftsin receptor on the surface of white cells which accumulate at the
CC site of infection and inflammation and can then be detected using
CC radiation detecting probes or by imaging using a planar or ring gamma
CC camera. The radiopharmaceuticals can also be used in treating cancer. The
CC present sequence represents the tuftsin receptor antagonist peptide given
CC in the present invention
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 29; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 1 TKPPR 5

RESULT 4
ABB08442
ID ABB08442 standard; peptide; 5 AA.
XX
AC ABB08442;
XX
DT 01-JUL-2002 (first entry)
XX
DE Tuftsin receptor antagonist (TKPPR).
XX
KW Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
KW angiogenesis; tumour cell; cytostatic; antagonist.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "residue optionally modified by one of the
FT following; Thr(Obzl), DPPE-Glutaroyl-
FT Di(aminodioxaoctanoyl), aminodioxaoctanoyl-Thr(OBnzl),
FT di(aminodioxaoctanoyl)-Thr(OBzl), (t-Bu), Fmoc-Thr(t-Bu),
FT H-Thr(t-Bu), Boc-Thr(But)"
FT Modified-site 2
FT /note= "residue optionally modified by one of the
FT following; Z (not further defined) or (Boc)"
FT Modified-site 5
FT /note= "residue optionally modified by one of the
FT following; OH, (NO2)Obzl, (Pmc)-t-Bu, (Pmc)-COHN"
XX
PN WO200191805-A2.
XX
PD 06-DEC-2001.
XX
PF 04-JUN-2001; 2001WO-US018053.
XX
PR 02-JUN-2000; 2000US-00585364.
XX
PA (BRAC) BRACCO RES USA.
XX
PI Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
PI Tweedle MF, Linder K, Nanjappan P, Raju N;
XX
DR WPI; 2002-195523/25.
XX
PT Composition for use in targeting endothelial cells, tumor cells or other
PT cells which express NP-1 comprises a compound containing a polypeptide,
PT linker and substrate.
XX
PS Claim 1; Page 2; 146pp; English.
XX
CC The invention relates to a composition for use in targeting endothelial
CC cells, tumour cells, or other cells which express NP-1. The activity of
CC compositions of the invention may be described as cytostatic. Compounds
CC of the invention are useful in pharmaceutical compositions for inhibiting
CC angiogenesis, for imaging and targeting an angiogenic site, endothelial
CC cells, tumour cells or other cells that express NP-1 in a human or
CC animal. They may also be used as ultrasound contrast agents, for staging
CC a tumour in a human or animal, for screening for the ability of an agent
CC to target endothelial cells, tumour cells or other cells that express NP-

```
CC   1. They may be used for therapeutic delivery in vivo of a bioactive agent
CC   or for treating an individual exhibiting effects of an angiogenesis or a
CC   related disorder. They may be used for delivering desired nucleic acids
CC   to endothelial cells, tumour cells or other cells expressing NP-1, for
CC   enhancing endothelial or tumour cell-targeted gene therapy, or gene
CC   therapy targeting angiogenic cells, and for treating a human or animal
CC   with a tumour or angiogenesis-related disease. The current sequence
CC   represents the tuftsin receptor antagonist (TKPPR), of which monomers,
CC   multimers, polymers or analogues of, may be used to target endothelial
CC   cells, or cells that express markers in common with endothelial cells
XX
SQ   Sequence 5 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 5
AAM51906
ID   AAM51906 standard; peptide; 5 AA.
XX
AC   AAM51906;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Tertiary ligand complex peptide #8.
XX
KW   Ternary ligand complex; highly functionalised phosphine ligand;
KW   disease site; cardiovascular disorder; thromboembolic disease;
KW   atherosclerosis; infectious disease; cancer; radiopharmaceutical.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "modified by 6-aminohexanamide"
XX
PN   WO200177122-A1.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-US011387.
XX
PR   07-APR-2000; 2000US-0195235P.
XX
PA   (DUPO ) DUPONT PHARM CO.
XX
PI   Liu S;
XX
DR   WPI; 2002-010884/01.
XX
PT   New ancillary ligands are useful as ligands for radiopharmaceuticals,
PT   which are useful for e.g. radioimaging a patient, diagnosing
PT   thromboembolic disorders, atherosclerosis, infections and inflammation.
XX
PS   Example; Page 126; 210pp; English.
XX
CC   The present invention relates to novel highly functionalised phosphine
CC   ligands which can be used as ancillary ligands in radiopharmaceuticals.
CC   These radiopharmaceuticals can then be used to radioimage a patient,
CC   enabling the visualisation of platelet deposition and disease sites, and
CC   allowing the diagnosis of infection, inflammation, transplant rejection,
CC   cardiovascular diseases such as thromboembolic disorders, and cancer. The
CC   present sequence is a peptide which formed part of a ligand in the
CC   exemplification of the invention
XX
SQ   Sequence 5 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 6
ADD10684
ID   ADD10684 standard; peptide; 5 AA.
XX
AC   ADD10684;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Tuftsin analogue peptide #1.
XX
KW   Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
KW   antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
KW   visualisation therapy; radiotherapy.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "Thr is optionally linked to an oregon green (OG)
FT                   moiety"
XX
PN   US2002147136-A1.
XX
PD   10-OCT-2002.
XX
PF   04-JUN-2001; 2001US-00871974.
XX
PR   02-JUN-2000; 2000US-00585364.
XX
PA   (VWRO/) VON WRONSKI M A.
PA   (MARI/) MARINELLI E R.
PA   (NUNN/) NUNN A D.
PA   (PILL/) PILLAI R.
PA   (RAMA/) RAMALINGAM K.
PA   (TWEE/) TWEEDLE M F.
PA   (LIND/) LINDER K.
PA   (NANJ/) NANJAPPAN P.
PA   (RAJU/) RAJU N.
XX
PI   Von Wronski MA,  Marinelli ER,  Nunn AD,  Pillai R,  Ramalingam K;
PI   Tweedle MF,  Linder K,  Nanjappan P,  Raju N;
XX
DR   WPI; 2003-800817/75.
XX
PT   Composition used in targeting endothelial cells e.g. tumor cells
PT   comprises compounds containing monomers, multimers or polymers of L-
PT   arginine-L-threonyl-L-lysyl-L-prolyl-L-prolyl.
XX
PS   Claim 1; SEQ ID NO 2; 85pp; English.
XX
CC   The invention relates to a composition (A1) comprising compounds
CC   containing monomers, multimers or polymers of TKPPR (ADD10684).
CC   Composition (A1) comprises a compound of formula A-L-B 1, where A is the
CC   TKPPR peptide, L is a linker moiety (of formula given in the
CC   specification) and B is a substrate (or a phospholipid group,
CC   derivatisable bead attached to a fluorescent or radioactive marker,
CC   bioactive agent, delivery vehicle for genetic material, drug or
CC   therapeutic, or chelating group (preferably N 4 , S 4 , N 3 S, N 2 S 2 or
CC   NS 3 ) comprising oxa-PnAO complexed with 99m Tc). The compound
CC   specifically binds to NP-1 (Vascular endothelial growth factor binding
CC   receptor transmembrane glycoprotein) or cells that express NP-1 with
CC   avidity of at least that of TKPPR. Also included are an ultrasound
```

CC contrast agent (c1) comprising a suspension of gas filled microbubbles
CC comprising the TKPPR compound, an ultrasound contrast agent (c2)
CC comprising a suspension of gas filled microballoons comprising the TKPPR
CC compound, preparation of the TKPPR compound (which comprises conjugating
CC the monomer, multimer or polymer of TKPPR or its analogue with a linker
CC to obtain a compound of formula A-L, forming a covalent or non-covalent
CC bond between A-L and the substrate B 1 or forming a covalent bond between
CC B 1 and the linker to form a conjugate B-L followed by conjugation with
CC the monomer), and a kit for preparing a radiopharmaceutical comprising
CC the compound. The compound used for targeting endothelial cells, tumour
CC cells or other cells which express NP-1, for inhibiting angiogenesis, for
CC ultrasound imaging, staging a tumour, screening at least one targeted
CC ultrasound contrast agent for the ability to target endothelial cells,
CC tumour cells or other cells which express NP-1, for the therapeutic
CC delivery in vivo of a bioactive agent and for delivering desired nucleic
CC acids to endothelial cells, tumour cells or other cells which express NP-
CC 1. The composition is also useful for visualisation therapy or
CC radiotherapy of endothelial cells. The present sequence is the TKPPR
CC peptide.
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 29; DB 7; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 1 TKPPR 5

RESULT 7
AAR85540
ID AAR85540 standard; peptide; 6 AA.
XX
AC AAR85540;
XX
DT 17-APR-1996 (first entry)
XX
DE Tuftsin analogue conjugated to metal-chelating moiety.
XX
KW tuftsin; chelator; radiodiagnostic; imaging; leukocytes; radionuclide.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "the amino group of this residue is acylated with
FT (S-Acm-mercaptoacetyl)-Ser- (N-methyl)hydrazino-
FT nicotinoyl"
XX
PN WO9503280-A1.
XX
PD 02-FEB-1995.
XX
PF 18-JUL-1994; 94WO-CA000395.
XX
PR 19-JUL-1993; 93US-00092911.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Pollak A, Kirby RA, Dufault R;
XX
DR WPI; 1995-075166/10.
XX
PT Hydrazino-type radionuclide chelating agents - capable of conjugation to
PT targetting molecules such as proteins, peptide(s) or antibodies and thus
PT useful in diagnosis and therapy.
XX
PS Claim 29; Page 29; 35pp; English.
XX
CC The patent provides radionuclide chelating compounds and their
CC conjugation products with targetting molecules such as proteins, peptides
CC or antibodies. The labelled targetting molecules may be used for in-vivo
CC diagnosis and therapy. The chelating compound is a hydrazino-containing,
CC N-heterocycle-containing compound, and the conjugated product is
CC exemplified by S-Acm-mercaptoacetyl-Ser-N- methyl-hydrazino-nicotinic
CC acid-Gly-Thr-Lys-Pro-Pro-Arg. The peptide portion of this molecule (the
CC present sequence) represents the tuftsin antagonist TKPPR condensed via
CC Gly onto the chelating moiety, giving a molecule which will chelate a
CC radionuclide such as 99mTc and target it towards leukocytes
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 2 TKPPR 6

RESULT 8
AAW11053
ID AAW11053 standard; peptide; 6 AA.
XX
AC AAW11053;
XX
DT 03-JUN-1997 (first entry)
XX
DE Leukocyte-targetting peptide used in diagnostic imaging.
XX
KW Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
KW diagnostic imaging.
XX
OS Synthetic.
XX
PN WO9603427-A1.
XX
PD 08-FEB-1996.
XX
PF 28-APR-1995; 95WO-CA000249.
XX
PR 22-JUL-1994; 94US-00279155.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Pollak A, Goodbody A;
XX
DR WPI; 1996-116994/12.
XX
PT New peptide derived radionuclide chelators and metal complexes - useful
PT for diagnostic imaging.
XX
PS Claim 13; Page 20; 30pp; English.
XX
CC AAW11053 is a peptide used for targetting agents to leukocytes. This
CC peptide can be coupled to a metal radionuclide chelator (structural
CC formula given in the specification), labelled with a diagnostically
CC useful metal isotope, to form a peptide derived radionuclide chelator.
CC When the chelator is coupled to a targeting molecule and labelled with a
CC diagnostically useful metal, it can be used to detect pathological
CC conditions by diagnostic imaging. For example, AAW11053 targets the
CC chelator to leukocytes and are useful for the rapid imaging of sites of
CC local inflammation. Radionuclides used include 99mTc, 64Cu, 67Cu, 97Ru,
CC 105Rh, 109Pd, 186Re, 188Re, 198Au, 199Au, 203Pb, 212Pb and 212Bi. The
CC coupling of a targeting agent and radionuclide using a chelating agent is
CC an alternative to the direct labelling of targeting agents in which
CC radionuclides are typically bound at the more numerous low- affinity
CC sites, forming unstable complexes. The new conjugates give better
CC scintigraphic images in rat inflammation studies than known imaging
CC agents Ga-67, 99mTc-IgG, 111In-WBC and 99mTc-Nanocoll. They image more
CC rapidly than the known agents and show superior biodistribution
XX
SQ Sequence 6 AA;

```
  Query Match             100.0%;  Score 29;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          2 TKPPR 6


RESULT 9
AAW31148
ID   AAW31148 standard; peptide; 6 AA.
XX
AC   AAW31148;
XX
DT   25-MAR-2003  (revised)
DT   23-JAN-1998  (first entry)
XX
DE   Amyloid plaque-targeting peptide-radionuclide chelator conjugate.
XX
KW   Target; delivery; radionuclide chelator; diagnosis; therapy; detection;
KW   atherosclerosis; thrombosis; Alzheimer's disease.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N terminally acetylated with S-Acm-Mercaptoacetyl
FT                   -Ser-N- methylhydrazino nicotinic acid"
XX
PN   US5659041-A.
XX
PD   19-AUG-1997.
XX
PF   02-SEP-1994;   94US-00299636.
XX
PR   19-JUL-1993;   93US-00092911.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Pollak A,  Kirby RA,  Dunn-Dufault R;
XX
DR   WPI; 1997-424290/39.
XX
PT   New thio:acetyl-aminoacid hydrazide compounds - useful as chemical
PT   chelator of radionuclides for radio:imaging of target tissues of
PT   diagnostic interest.
XX
PS   Example 11; Col 14; 20pp; English.
XX
CC   AAW31148 is a peptide conjugate of an amyloid plaque-targeting peptide
CC   and a new hydrazino-type compound. The new compound is a radionuclide
CC   chelator and is used to radiolabel targeting peptides for the detection
CC   and diagnostic imaging of sites of disease, e.g. amyloid plaques in
CC   Alzheimer's disease, thrombi in thrombosis or other sites of infection.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 6 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          2 TKPPR 6


RESULT 10
AAW31147
ID   AAW31147 standard; peptide; 6 AA.
XX
AC   AAW31147;
XX
DT   25-MAR-2003  (revised)
DT   23-JAN-1998  (first entry)
XX
DE   Amyloid plaque-targeting peptide.
XX
KW   Target; delivery; radionuclide chelator; diagnosis; therapy; detection;
KW   atherosclerosis; thrombosis; Alzheimer's disease.
XX
OS   Synthetic.
XX
PN   US5659041-A.
XX
PD   19-AUG-1997.
XX
PF   02-SEP-1994;   94US-00299636.
XX
PR   19-JUL-1993;   93US-00092911.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Pollak A,  Kirby RA,  Dunn-Dufault R;
XX
DR   WPI; 1997-424290/39.
XX
PT   New thio:acetyl-aminoacid hydrazide compounds - useful as chemical
PT   chelator of radionuclides for radio:imaging of target tissues of
PT   diagnostic interest.
XX
PS   Disclosure; Col 29; 20pp; English.
XX
CC   AAW31110-W31147 are peptides used for targeting a new hydrazino-type
CC   compound to various sites of disease, e.g. atherosclerotic plaque, sites
CC   of infection, platelets, thrombus or amyloid plaque. The new compound is
CC   a radionuclide chelator and is used to radiolabel the targeting peptides
CC   for the detection and diagnostic imaging of sites of disease, e.g.
CC   amyloid plaques in Alzheimer's disease. (Updated on 25-MAR-2003 to
CC   correct PF field.)
XX
SQ   Sequence 6 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          2 TKPPR 6


RESULT 11
AAY49840
ID   AAY49840 standard; peptide; 6 AA.
XX
AC   AAY49840;
XX
DT   20-JAN-2000  (first entry)
XX
DE   Tuftsin receptor antagonist chelate conjugate #1.
XX
KW   Tuftsin receptor antagonist; chelate conjugate; radiopharmaceutical;
KW   diagnosis; infection; inflammation; imaging; cancer; tumour.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= Acp
FT                   /note= "modified 6-aminocaproic acid: 6-((6-((1-aza-2- (2
FT                   -sulphophenyl)vinyl)amino)-3-pyridyl)
FT                   carbonylamino)hexanoyl"
XX
```

PN WO9951628-A1.
XX
PD 14-OCT-1999.
XX
PF 29-MAR-1999; 99WO-US006824.
XX
PR 03-APR-1998; 98US-0080672P.
XX
PA (DUPO) DU PONT PHARM CO.
XX
PI Edwards DS, Rajopadhye M;
XX
DR WPI; 1999-633729/54.
XX
PT New polypeptides and radiopharmacueticals used for imaging infection,
PT inflammation and cancer.
XX
PS Claim 4; Page 71; 80pp; English.
XX
CC The present invention describes polypeptide compounds of formula (I),
CC capable of direct transformation into a radiopharmaceutical: Ch-Ln-(X1-X2
CC -X3-X4-X5)d (I), where X1-X5 = amino acids; Ln = a linking group; Ch = a
CC metal bonding unit; and d is selected from 1, 2 and 3. The
CC radiopharmaceuticals are useful for the diagnosis of infection,
CC inflammation and cancer. The radiopharmaceuticals bind in vivo to the
CC tuftsin receptor on the surface of white cells which accumulate at the
CC site of infection and inflammation and can then be detected using
CC radiation detecting probes or by imaging using a planar or ring gamma
CC camera. The radiopharmaceuticals can also be used in treating cancer. The
CC present sequence represents a specifically claimed tuftsin receptor
CC antagonist chelate conjugate from the present invention
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 2 TKPPR 6

RESULT 12
ABB08444
ID ABB08444 standard; peptide; 6 AA.
XX
AC ABB08444;
XX
DT 01-JUL-2002 (first entry)
XX
DE Tuftsin receptor antagonist (TKPPR) derivative peptide 2.
XX
KW Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
KW angiogenesis; tumour cell; cytostatic; antagonist.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "residue optionally modified by one of the
FT following; DPPE-Glutaroyl, N-Glutaroyl-Gly, NH2, Oregon
FT Green (5 isomer), Fmoc"
FT Modified-site 2
FT /note= "residue optionally modified by one of the
FT following; Obzl, (tBu)"
FT Modified-site 3
FT /note= "residue optionally modified by one of the
FT following; Z (not further defined), (Mtt), (Boc)"
FT Modified-site 6
FT /note= "residue optionally modified by one of the
FT following; OH, (No2)Obzl, (NO2), (Pmc)OtBu, (pmc)-Wang-
FT resin"
XX
PN WO200191805-A2.
XX
PD 06-DEC-2001.
XX
PF 04-JUN-2001; 2001WO-US018053.
XX
PR 02-JUN-2000; 2000US-00585364.
XX
PA (BRAC) BRACCO RES USA.
XX
PI Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
PI Tweedle MF, Linder K, Nanjappan P, Raju N;
XX
DR WPI; 2002-195523/25.
XX
PT Composition for use in targeting endothelial cells, tumor cells or other
PT cells which express NP-1 comprises a compound containing a polypeptide,
PT linker and substrate.
XX
PS Example 4; Page 67; 146pp; English.
XX
CC The invention relates to a composition for use in targeting endothelial
CC cells, tumour cells, or other cells which express NP-1. The activity of
CC compositions of the invention may be described as cytostatic. Compounds
CC of the invention are useful in pharmaceutical compositions for inhibiting
CC angiogenesis, for imaging and targeting an angiogenic site, endothelial
CC cells, tumour cells or other cells that express NP-1 in a human or
CC animal. They may also be used as ultrasound contrast agents, for staging
CC a tumour in a human or animal, for screening for the ability of an agent
CC to target endothelial cells, tumour cells or other cells that express NP-
CC 1. They may be used for therapeutic delivery in vivo of a bioactive agent
CC or for treating an individual exhibiting effects of an angiogenesis or a
CC related disorder. They may be used for delivering desired nucleic acids
CC to endothelial cells, tumour cells or other cells expressing NP-1, for
CC enhancing endothelial or tumour cell-targeted gene therapy, or gene
CC therapy targeting angiogenic cells, and for treating a human or animal
CC with a tumour or angiogenesis-related disease. The current sequence
CC represents a tuftsin receptor antagonist (TKPPR) derivative of the
CC invention
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 2 TKPPR 6

RESULT 13
AAM51904
ID AAM51904 standard; peptide; 6 AA.
XX
AC AAM51904;
XX
DT 01-FEB-2002 (first entry)
XX
DE Tertiary ligand complex peptide #6.
XX
KW Ternary ligand complex; highly functionalised phosphine ligand;
KW disease site; cardiovascular disorder; thromboembolic disease;
KW atherosclerosis; infectious disease; cancer; radiopharmaceutical.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /note= "D-form residue"
FT Modified-site 2
FT /label= OTHER

```
FT                  /note= "modified by 6-aminohexanamide"
XX
PN   WO200177122-A1.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-US011387.
XX
PR   07-APR-2000; 2000US-0195235P.
XX
PA   (DUPO ) DUPONT PHARM CO.
XX
PI   Liu S;
XX
DR   WPI; 2002-010884/01.
XX
PT   New ancillary ligands are useful as ligands for radiopharmaceuticals,
PT   which are useful for e.g. radioimaging a patient, diagnosing
PT   thromboembolic disorders, atherosclerosis, infections and inflammation.
XX
PS   Example; Page 125; 210pp; English.
XX
CC   The present invention relates to novel highly functionalised phosphine
CC   ligands which can be used as ancillary ligands in radiopharmaceuticals.
CC   These radiopharmaceuticals can then be used to radioimage a patient,
CC   enabling the visualisation of platelet deposition and disease sites, and
CC   allowing the diagnosis of infection, inflammation, transplant rejection,
CC   cardiovascular diseases such as thromboembolic disorders, and cancer. The
CC   present sequence is a peptide which formed part of a ligand in the
CC   exemplification of the invention
XX
SQ   Sequence 6 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          2 TKPPR 6


RESULT 14
AAM51905
ID   AAM51905 standard; peptide; 6 AA.
XX
AC   AAM51905;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Tertiary ligand complex peptide #7.
XX
KW   Ternary ligand complex; highly functionalised phosphine ligand;
KW   disease site; cardiovascular disorder; thromboembolic disease;
KW   atherosclerosis; infectious disease; cancer; radiopharmaceutical.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Modified-site   2
FT                   /label= OTHER
FT                   /note= "modified by 6-aminohexanamide"
XX
PN   WO200177122-A1.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-US011387.
XX
PR   07-APR-2000; 2000US-0195235P.
XX
PA   (DUPO ) DUPONT PHARM CO.
XX
PI   Liu S;
XX
DR   WPI; 2002-010884/01.
XX
PT   New ancillary ligands are useful as ligands for radiopharmaceuticals,
PT   which are useful for e.g. radioimaging a patient, diagnosing
PT   thromboembolic disorders, atherosclerosis, infections and inflammation.
XX
PS   Example; Page 125; 210pp; English.
XX
CC   The present invention relates to novel highly functionalised phosphine
CC   ligands which can be used as ancillary ligands in radiopharmaceuticals.
CC   These radiopharmaceuticals can then be used to radioimage a patient,
CC   enabling the visualisation of platelet deposition and disease sites, and
CC   allowing the diagnosis of infection, inflammation, transplant rejection,
CC   cardiovascular diseases such as thromboembolic disorders, and cancer. The
CC   present sequence is a peptide which formed part of a ligand in the
CC   exemplification of the invention
XX
SQ   Sequence 6 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          2 TKPPR 6


RESULT 15
ADD10689
ID   ADD10689 standard; peptide; 6 AA.
XX
AC   ADD10689;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Tuftsin analogue peptide #4.
XX
KW   Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
KW   antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
KW   visualisation therapy; radiotherapy.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "Glu is covalently linked to a DPPE
FT                   (dipalmitoylphosphatidylethanolamine) moiety"
FT   Modified-site   2
FT                   /label= OTHER
FT                   /note= "Thr is di(aminodioxaoctanoyl)-Thr"
XX
PN   US2002147136-A1.
XX
PD   10-OCT-2002.
XX
PF   04-JUN-2001; 2001US-00871974.
XX
PR   02-JUN-2000; 2000US-00585364.
XX
PA   (VWRO/) VON WRONSKI M A.
PA   (MARI/) MARINELLI E R.
PA   (NUNN/) NUNN A D.
PA   (PILL/) PILLAI R.
PA   (RAMA/) RAMALINGAM K.
PA   (TWEE/) TWEEDLE M F.
PA   (LIND/) LINDER K.
PA   (NANJ/) NANJAPPAN P.
```

PA (RAJU/) RAJU N.
XX
PI Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
PI Tweedle MF, Linder K, Nanjappan P, Raju N;
XX
DR WPI; 2003-800817/75.
XX
PT Composition used in targeting endothelial cells e.g. tumor cells
PT comprises compounds containing monomers, multimers or polymers of L-
PT arginine-L-threonyl-L-lysyl-L-prolyl-L-prolyl.
XX
PS Example 23; Page 44; 85pp; English.
XX
CC The invention relates to a composition (A1) comprising compounds
CC containing monomers, multimers or polymers of TKPPR (ADD10684).
CC Composition (A1) comprises a compound of formula A-L-B 1, where A is the
CC TKPPR peptide, L is a linker moiety (of formula given in the
CC specification) and B is a substrate (or a phospholipid group,
CC derivatisable bead attached to a fluorescent or radioactive marker,
CC bioactive agent, delivery vehicle for genetic material, drug or
CC therapeutic, or chelating group (preferably N 4 , S 4 , N 3 S, N 2 S 2 or
CC NS 3) comprising oxa-PnAO complexed with 99m Tc). The compound
CC specifically binds to NP-1 (Vascular endothelial growth factor binding
CC receptor transmembrane glycoprotein) or cells that express NP-1 with
CC avidity of at least that of TKPPR. Also included are an ultrasound
CC contrast agent (c1) comprising a suspension of gas filled microbubbles
CC comprising the TKPPR compound, an ultrasound contrast agent (c2)
CC comprising a suspension of gas filled microballoons comprising the TKPPR
CC compound, preparation of the TKPPR compound (which comprises conjugating
CC the monomer, multimer or polymer of TKPPR or its analogue with a linker
CC to obtain a compound of formula A-L, forming a covalent or non-covalent
CC bond between A-L and the substrate B 1 or forming a covalent bond between
CC B 1 and the linker to form a conjugate B-L followed by conjugation with
CC the monomer), and a kit for preparing a radiopharmaceutical comprising
CC the compound. The compound used for targeting endothelial cells, tumour
CC cells or other cells which express NP-1, for inhibiting angiogenesis, for
CC ultrasound imaging, staging a tumour, screening at least one targeted
CC ultrasound contrast agent for the ability to target endothelial cells,
CC tumour cells or other cells which express NP-1, for the therapeutic
CC delivery in vivo of a bioactive agent and for delivering desired nucleic
CC acids to endothelial cells, tumour cells or other cells which express NP-
CC 1. The composition is also useful for visualisation therapy or
CC radiotherapy of endothelial cells. The present sequence is a TKPPR
CC analogue peptide.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 7; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 2 TKPPR 6

RESULT 16
ADD10687
ID ADD10687 standard; peptide; 6 AA.
XX
AC ADD10687;
XX
DT 01-JAN-2004 (first entry)
XX
DE Tuftsin analogue peptide #3.
XX
KW Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
KW antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
KW visualisation therapy; radiotherapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "Gly is covalently linked to a DPPE
FT (dipalmitoylphosphatidylethanolamine) moiety"
XX
PN US2002147136-A1.
XX
PD 10-OCT-2002.
XX
PF 04-JUN-2001; 2001US-00871974.
XX
PR 02-JUN-2000; 2000US-00585364.
XX
PA (VWRO/) VON WRONSKI M A.
PA (MARI/) MARINELLI E R.
PA (NUNN/) NUNN A D.
PA (PILL/) PILLAI R.
PA (RAMA/) RAMALINGAM K.
PA (TWEE/) TWEEDLE M F.
PA (LIND/) LINDER K.
PA (NANJ/) NANJAPPAN P.
PA (RAJU/) RAJU N.
XX
PI Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
PI Tweedle MF, Linder K, Nanjappan P, Raju N;
XX
DR WPI; 2003-800817/75.
XX
PT Composition used in targeting endothelial cells e.g. tumor cells
PT comprises compounds containing monomers, multimers or polymers of L-
PT arginine-L-threonyl-L-lysyl-L-prolyl-L-prolyl.
XX
PS Example 14; Page 38; 85pp; English.
XX
CC The invention relates to a composition (A1) comprising compounds
CC containing monomers, multimers or polymers of TKPPR (ADD10684).
CC Composition (A1) comprises a compound of formula A-L-B 1, where A is the
CC TKPPR peptide, L is a linker moiety (of formula given in the
CC specification) and B is a substrate (or a phospholipid group,
CC derivatisable bead attached to a fluorescent or radioactive marker,
CC bioactive agent, delivery vehicle for genetic material, drug or
CC therapeutic, or chelating group (preferably N 4 , S 4 , N 3 S, N 2 S 2 or
CC NS 3) comprising oxa-PnAO complexed with 99m Tc). The compound
CC specifically binds to NP-1 (Vascular endothelial growth factor binding
CC receptor transmembrane glycoprotein) or cells that express NP-1 with
CC avidity of at least that of TKPPR. Also included are an ultrasound
CC contrast agent (c1) comprising a suspension of gas filled microbubbles
CC comprising the TKPPR compound, an ultrasound contrast agent (c2)
CC comprising a suspension of gas filled microballoons comprising the TKPPR
CC compound, preparation of the TKPPR compound (which comprises conjugating
CC the monomer, multimer or polymer of TKPPR or its analogue with a linker
CC to obtain a compound of formula A-L, forming a covalent or non-covalent
CC bond between A-L and the substrate B 1 or forming a covalent bond between
CC B 1 and the linker to form a conjugate B-L followed by conjugation with
CC the monomer), and a kit for preparing a radiopharmaceutical comprising
CC the compound. The compound used for targeting endothelial cells, tumour
CC cells or other cells which express NP-1, for inhibiting angiogenesis, for
CC ultrasound imaging, staging a tumour, screening at least one targeted
CC ultrasound contrast agent for the ability to target endothelial cells,
CC tumour cells or other cells which express NP-1, for the therapeutic
CC delivery in vivo of a bioactive agent and for delivering desired nucleic
CC acids to endothelial cells, tumour cells or other cells which express NP-
CC 1. The composition is also useful for visualisation therapy or
CC radiotherapy of endothelial cells. The present sequence is a TKPPR
CC analogue peptide.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 7; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5

```
                  |||||
Db          2 TKPPR 6


RESULT 17
AAY49842
ID   AAY49842 standard; peptide; 7 AA.
XX
AC   AAY49842;
XX
DT   20-JAN-2000  (first entry)
XX
DE   Tuftsin receptor antagonist chelate conjugate #3.
XX
KW   Tuftsin receptor antagonist; chelate conjugate; radiopharmaceutical;
KW   diagnosis; infection; inflammation; imaging; cancer; tumour.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "modified tyrosine: ((6-((1-aza-2- (2-
FT                   sulphophenyl)vinyl)amino)-3-pyridyl) carbonyl)-L-
FT                   tyrosinyl"
FT   Modified-site   2
FT                   /label= Acp
XX
PN   WO9951628-A1.
XX
PD   14-OCT-1999.
XX
PF   29-MAR-1999;   99WO-US006824.
XX
PR   03-APR-1998;   98US-0080672P.
XX
PA   (DUPO ) DU PONT PHARM CO.
XX
PI   Edwards DS,  Rajopadhye M;
XX
DR   WPI; 1999-633729/54.
XX
PT   New polypeptides and radiopharmacueticals used for imaging infection,
PT   inflammation and cancer.
XX
PS   Claim 4; Page 71; 80pp; English.
XX
CC   The present invention describes polypeptide compounds of formula (I),
CC   capable of direct transformation into a radiopharmaceutical: Ch-Ln-(X1-X2
CC   -X3-X4-X5)d (I), where X1-X5 = amino acids; Ln = a linking group; Ch = a
CC   metal bonding unit; and d is selected from 1, 2 and 3. The
CC   radiopharmaceuticals are useful for the diagnosis of infection,
CC   inflammation and cancer. The radiopharmaceuticals bind in vivo to the
CC   tuftsin receptor on the surface of white cells which accumulate at the
CC   site of infection and inflammation and can then be detected using
CC   radiation detecting probes or by imaging using a planar or ring gamma
CC   camera. The radiopharmaceuticals can also be used in treating cancer. The
CC   present sequence represents a specifically claimed tuftsin receptor
CC   antagonist chelate conjugate from the present invention
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          3 TKPPR 7


RESULT 18
AAY49841
ID   AAY49841 standard; peptide; 7 AA.
XX
AC   AAY49841;
XX
DT   20-JAN-2000  (first entry)
XX
DE   Tuftsin receptor antagonist chelate conjugate #2.
XX
KW   Tuftsin receptor antagonist; chelate conjugate; radiopharmaceutical;
KW   diagnosis; infection; inflammation; imaging; cancer; tumour.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "modified phenylalanine: ((6-((1-aza-2- (2-
FT                   sulphophenyl)vinyl)amino)-3-pyridyl) carbonyl)-L-
FT                   phenylalanyl"
FT   Modified-site   2
FT                   /label= Acp
XX
PN   WO9951628-A1.
XX
PD   14-OCT-1999.
XX
PF   29-MAR-1999;   99WO-US006824.
XX
PR   03-APR-1998;   98US-0080672P.
XX
PA   (DUPO ) DU PONT PHARM CO.
XX
PI   Edwards DS,  Rajopadhye M;
XX
DR   WPI; 1999-633729/54.
XX
PT   New polypeptides and radiopharmacueticals used for imaging infection,
PT   inflammation and cancer.
XX
PS   Claim 4; Page 71; 80pp; English.
XX
CC   The present invention describes polypeptide compounds of formula (I),
CC   capable of direct transformation into a radiopharmaceutical: Ch-Ln-(X1-X2
CC   -X3-X4-X5)d (I), where X1-X5 = amino acids; Ln = a linking group; Ch = a
CC   metal bonding unit; and d is selected from 1, 2 and 3. The
CC   radiopharmaceuticals are useful for the diagnosis of infection,
CC   inflammation and cancer. The radiopharmaceuticals bind in vivo to the
CC   tuftsin receptor on the surface of white cells which accumulate at the
CC   site of infection and inflammation and can then be detected using
CC   radiation detecting probes or by imaging using a planar or ring gamma
CC   camera. The radiopharmaceuticals can also be used in treating cancer. The
CC   present sequence represents a specifically claimed tuftsin receptor
CC   antagonist chelate conjugate from the present invention
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          3 TKPPR 7


RESULT 19
ABB08447
ID   ABB08447 standard; peptide; 7 AA.
XX
AC   ABB08447;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Tuftsin receptor antagonist (TKPPR) derivative dimer.
XX
```

```
KW   Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
KW   angiogenesis; tumour cell; cytostatic; antagonist.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   6
FT                   /note= "residue modified by the addition of the following
FT                   block of residues (RPPKT)"
FT   Modified-site   7
FT                   /note= "residue modified by TTDA-OG (TTDA= 4,7,10-Trioxa-
FT                   1,13-TridecaneDiamine, OG= Oregon green)"
XX
PN   WO200191805-A2.
XX
PD   06-DEC-2001.
XX
PF   04-JUN-2001; 2001WO-US018053.
XX
PR   02-JUN-2000; 2000US-00585364.
XX
PA   (BRAC ) BRACCO RES USA.
XX
PI   Von Wronski MA,  Marinelli ER,  Nunn AD,  Pillai R,  Ramalingam K;
PI   Tweedle MF,  Linder K,  Nanjappan P,  Raju N;
XX
DR   WPI; 2002-195523/25.
XX
PT   Composition for use in targeting endothelial cells, tumor cells or other
PT   cells which express NP-1 comprises a compound containing a polypeptide,
PT   linker and substrate.
XX
PS   Example 29; Page 113; 146pp; English.
XX
CC   The invention relates to a composition for use in targeting endothelial
CC   cells, tumour cells, or other cells which express NP-1. The activity of
CC   compositions of the invention may be described as cytostatic. Compounds
CC   of the invention are useful in pharmaceutical compositions for inhibiting
CC   angiogenesis, for imaging and targeting an angiogenic site, endothelial
CC   cells, tumour cells or other cells that express NP-1 in a human or
CC   animal. They may also be used as ultrasound contrast agents, for staging
CC   a tumour in a human or animal, for screening for the ability of an agent
CC   to target endothelial cells, tumour cells or other cells that express NP-
CC   1. They may be used for therapeutic delivery in vivo of a bioactive agent
CC   or for treating an individual exhibiting effects of an angiogenesis or a
CC   related disorder. They may be used for delivering desired nucleic acids
CC   to endothelial cells, tumour cells or other cells expressing NP-1, for
CC   enhancing endothelial or tumour cell-targeted gene therapy, or gene
CC   therapy targeting angiogenic cells, and for treating a human or animal
CC   with a tumour or angiogenesis-related disease. The current sequence
CC   represents a tuftsin receptor antagonist (TKPPR) derivative dimer of the
CC   invention
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 20
ABB08445
ID   ABB08445 standard; peptide; 7 AA.
XX
AC   ABB08445;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Linear and cyclic CTKPPR (BRU-305 and BRU-306).
XX
KW   Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
KW   angiogenesis; tumour cell; cytostatic; antagonist.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Disufide_bond   1. .7
FT                   /note= "optionally present"
FT   Modified-site   1
FT                   /note= "Fmoc-Cys(Trt)-OH, residue optionally modified by
FT                   H group"
FT   Modified-site   2
FT                   /note= "Fmoc-Thr(But)-OH"
FT   Modified-site   3
FT                   /note= "Fmoc-Lys(Boc)-OH"
FT   Modified-site   4
FT                   /note= "Fmoc-Pro-OH"
FT   Modified-site   5
FT                   /note= "Fmoc-Pro-OH"
FT   Modified-site   6
FT                   /note= "Fmoc-Arg(Pmc)-OH"
FT   Modified-site   7
FT                   /note= "Fmoc-Cys(Trt)-OH; residue optionally modified by
FT                   an OH group"
XX
PN   WO200191805-A2.
XX
PD   06-DEC-2001.
XX
PF   04-JUN-2001; 2001WO-US018053.
XX
PR   02-JUN-2000; 2000US-00585364.
XX
PA   (BRAC ) BRACCO RES USA.
XX
PI   Von Wronski MA,  Marinelli ER,  Nunn AD,  Pillai R,  Ramalingam K;
PI   Tweedle MF,  Linder K,  Nanjappan P,  Raju N;
XX
DR   WPI; 2002-195523/25.
XX
PT   Composition for use in targeting endothelial cells, tumor cells or other
PT   cells which express NP-1 comprises a compound containing a polypeptide,
PT   linker and substrate.
XX
PS   Example 21; Page 81; 146pp; English.
XX
CC   The invention relates to a composition for use in targeting endothelial
CC   cells, tumour cells, or other cells which express NP-1. The activity of
CC   compositions of the invention may be described as cytostatic. Compounds
CC   of the invention are useful in pharmaceutical compositions for inhibiting
CC   angiogenesis, for imaging and targeting an angiogenic site, endothelial
CC   cells, tumour cells or other cells that express NP-1 in a human or
CC   animal. They may also be used as ultrasound contrast agents, for staging
CC   a tumour in a human or animal, for screening for the ability of an agent
CC   to target endothelial cells, tumour cells or other cells that express NP-
CC   1. They may be used for therapeutic delivery in vivo of a bioactive agent
CC   or for treating an individual exhibiting effects of an angiogenesis or a
CC   related disorder. They may be used for delivering desired nucleic acids
CC   to endothelial cells, tumour cells or other cells expressing NP-1, for
CC   enhancing endothelial or tumour cell-targeted gene therapy, or gene
CC   therapy targeting angiogenic cells, and for treating a human or animal
CC   with a tumour or angiogenesis-related disease. The current sequence
CC   represents a tuftsin receptor antagonist (TKPPR) derivative of the
CC   invention referred to as linear and cyclic CTKPPR (BRU-305 and BRU-306)
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
```

```
                  |||||
Db          2 TKPPR 6


RESULT 21
ADD10688
ID   ADD10688 standard; peptide; 7 AA.
XX
AC   ADD10688;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Cyclic Tuftsin analogue peptide.
XX
KW   Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
KW   antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
KW   visualisation therapy; radiotherapy; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Disulfide-bond  1. .7
FT                   /note= "The molecule is cyclised via this disulfide bond"
XX
PN   US2002147136-A1.
XX
PD   10-OCT-2002.
XX
PF   04-JUN-2001; 2001US-00871974.
XX
PR   02-JUN-2000; 2000US-00585364.
XX
PA   (VWRO/) VON WRONSKI M A.
PA   (MARI/) MARINELLI E R.
PA   (NUNN/) NUNN A D.
PA   (PILL/) PILLAI R.
PA   (RAMA/) RAMALINGAM K.
PA   (TWEE/) TWEEDLE M F.
PA   (LIND/) LINDER K.
PA   (NANJ/) NANJAPPAN P.
PA   (RAJU/) RAJU N.
XX
PI   Von Wronski MA,  Marinelli ER,  Nunn AD,  Pillai R,  Ramalingam K;
PI   Tweedle MF,  Linder K,  Nanjappan P,  Raju N;
XX
DR   WPI; 2003-800817/75.
XX
PT   Composition used in targeting endothelial cells e.g. tumor cells
PT   comprises compounds containing monomers, multimers or polymers of L-
PT   arginine-L-threonyl-L-lysyl-L-prolyl-L-prolyl.
XX
PS   Example 21; Page 41; 85pp; English.
XX
CC   The invention relates to a composition (A1) comprising compounds
CC   containing monomers, multimers or polymers of TKPPR (ADD10684).
CC   Composition (A1) comprises a compound of formula A-L-B 1, where A is the
CC   TKPPR peptide, L is a linker moiety (of formula given in the
CC   specification) and B is a substrate (or a phospholipid group,
CC   derivatisable bead attached to a fluorescent or radioactive marker,
CC   bioactive agent, delivery vehicle for genetic material, drug or
CC   therapeutic, or chelating group (preferably N 4 , S 4 , N 3 S, N 2 S 2 or
CC   NS 3 ) comprising oxa-PnAO complexed with 99m Tc). The compound
CC   specifically binds to NP-1 (Vascular endothelial growth factor binding
CC   receptor transmembrane glycoprotein) or cells that express NP-1 with
CC   avidity of at least that of TKPPR. Also included are an ultrasound
CC   contrast agent (c1) comprising a suspension of gas filled microbubbles
CC   comprising the TKPPR compound, an ultrasound contrast agent (c2)
CC   comprising a suspension of gas filled microballoons comprising the TKPPR
CC   compound, preparation of the TKPPR compound (which comprises conjugating
CC   the monomer, multimer or polymer of TKPPR or its analogue with a linker
CC   to obtain a compound of formula A-L, forming a covalent or non-covalent
CC   bond between A-L and the substrate B 1 or forming a covalent bond between
CC   B 1 and the linker to form a conjugate B-L followed by conjugation with
CC   the monomer), and a kit for preparing a radiopharmaceutical comprising
CC   the compound. The compound used for targeting endothelial cells, tumour
CC   cells or other cells which express NP-1, for inhibiting angiogenesis, for
CC   ultrasound imaging, staging a tumour, screening at least one targeted
CC   ultrasound contrast agent for the ability to target endothelial cells,
CC   tumour cells or other cells which express NP-1, for the therapeutic
CC   delivery in vivo of a bioactive agent and for delivering desired nucleic
CC   acids to endothelial cells, tumour cells or other cells which express NP-
CC   1. The composition is also useful for visualisation therapy or
CC   radiotherapy of endothelial cells. The present sequence is a cyclic TKPPR
CC   analogue peptide.
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 29;  DB 7;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
                |||||
Db          2 TKPPR 6


RESULT 22
ADD10686
ID   ADD10686 standard; peptide; 7 AA.
XX
AC   ADD10686;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Tuftsin analogue peptide #2.
XX
KW   Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
KW   antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
KW   visualisation therapy; radiotherapy.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "Glu is covalently linked to a DPPE
FT                   (dipalmitoylphosphatidylethanolamine) moiety"
XX
PN   US2002147136-A1.
XX
PD   10-OCT-2002.
XX
PF   04-JUN-2001; 2001US-00871974.
XX
PR   02-JUN-2000; 2000US-00585364.
XX
PA   (VWRO/) VON WRONSKI M A.
PA   (MARI/) MARINELLI E R.
PA   (NUNN/) NUNN A D.
PA   (PILL/) PILLAI R.
PA   (RAMA/) RAMALINGAM K.
PA   (TWEE/) TWEEDLE M F.
PA   (LIND/) LINDER K.
PA   (NANJ/) NANJAPPAN P.
PA   (RAJU/) RAJU N.
XX
PI   Von Wronski MA,  Marinelli ER,  Nunn AD,  Pillai R,  Ramalingam K;
PI   Tweedle MF,  Linder K,  Nanjappan P,  Raju N;
XX
DR   WPI; 2003-800817/75.
XX
PT   Composition used in targeting endothelial cells e.g. tumor cells
PT   comprises compounds containing monomers, multimers or polymers of L-
PT   arginine-L-threonyl-L-lysyl-L-prolyl-L-prolyl.
XX
PS   Example 4; Page 34; 85pp; English.
```

```
XX
CC   The invention relates to a composition (A1) comprising compounds
CC   containing monomers, multimers or polymers of TKPPR (ADD10684).
CC   Composition (A1) comprises a compound of formula A-L-B 1, where A is the
CC   TKPPR peptide, L is a linker moiety (of formula given in the
CC   specification) and B is a substrate (or a phospholipid group,
CC   derivatisable bead attached to a fluorescent or radioactive marker,
CC   bioactive agent, delivery vehicle for genetic material, drug or
CC   therapeutic, or chelating group (preferably N 4 , S 4 , N 3 S, N 2 S 2 or
CC   NS 3 ) comprising oxa-PnAO complexed with 99m Tc). The compound
CC   specifically binds to NP-1 (Vascular endothelial growth factor binding
CC   receptor transmembrane glycoprotein) or cells that express NP-1 with
CC   avidity of at least that of TKPPR. Also included are an ultrasound
CC   contrast agent (c1) comprising a suspension of gas filled microbubbles
CC   comprising the TKPPR compound, an ultrasound contrast agent (c2)
CC   comprising a suspension of gas filled microballoons comprising the TKPPR
CC   compound, preparation of the TKPPR compound (which comprises conjugating
CC   the monomer, multimer or polymer of TKPPR or its analogue with a linker
CC   to obtain a compound of formula A-L, forming a covalent or non-covalent
CC   bond between A-L and the substrate B 1 or forming a covalent bond between
CC   B 1 and the linker to form a conjugate B-L followed by conjugation with
CC   the monomer), and a kit for preparing a radiopharmaceutical comprising
CC   the compound. The compound used for targeting endothelial cells, tumour
CC   cells or other cells which express NP-1, for inhibiting angiogenesis, for
CC   ultrasound imaging, staging a tumour, screening at least one targeted
CC   ultrasound contrast agent for the ability to target endothelial cells,
CC   tumour cells or other cells which express NP-1, for the therapeutic
CC   delivery in vivo of a bioactive agent and for delivering desired nucleic
CC   acids to endothelial cells, tumour cells or other cells which express NP-
CC   1. The composition is also useful for visualisation therapy or
CC   radiotherapy of endothelial cells. The present sequence is a TKPPR
CC   analogue peptide.
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 29;  DB 7;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          3 TKPPR 7


RESULT 23
AAR88740
ID   AAR88740 standard; peptide; 8 AA.
XX
AC   AAR88740;
XX
DT   10-APR-1996  (first entry)
XX
DE   Tuftsin antagonist peptide-metal chelator conjugate.
XX
KW   Peptide-chelator conjugate; metal chelator; diagnostic imaging;
KW   inflammation; radionuclide; tuftsin; analogue; antagonist.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .5
FT                   /label= tuftsin_antagonist
FT   Modified-site   6
FT                   /label= bAla
FT                   /note= "linking group"
FT   Modified-site   7
FT                   /label= OTHER
FT                   /note= "the side-chain (epsilon) amino group of Lys at
FT                   position 7 forms a peptide bond with the C-terminus of
FT                   the tetrapeptide N',N-dimethylglycyl-Ser-Cys(Acm)-Gly-;
FT                   this sidechain group consists of a tripeptide metal
FT                   chelator and a linking residue (Gly)"
XX
PN   WO9522996-A2.
XX
PD   31-AUG-1995.
XX
PF   24-FEB-1995;   95WO-CA000106.
XX
PR   25-FEB-1994;   94US-00202178.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Goodbody A,  Pollak A;
XX
DR   WPI; 1995-311386/40.
XX
PT   New peptide-chelator conjugate and complex with traceable metal - used to
PT   image sites of inflammation in vivo without significant accumulation on
PT   the gastrointestinal tract.
XX
PS   Claim 19; Page 21; 23pp; English.
XX
CC   The present sequence is that of a specifically claimed peptide- chelator
CC   conjugate in which a tuftsin antagonist peptide is coupled to a metal
CC   chelator, via a linking group. The chelator serves as a labelling site
CC   for radionuclide metals such as technetium-99m. The tuftsin antagonist
CC   targets the conjugate to macrophages and neutrophils at sites of
CC   inflammation without significant accumulation in the gastrointestinal
CC   tract (unlike the native tuftsin tetrapeptide). The conjugate is thus
CC   useful for diagnostic imaging of inflammation sites, providing an
CC   improved target to background ratio
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 24
AAR76218
ID   AAR76218 standard; peptide; 8 AA.
XX
AC   AAR76218;
XX
DT   12-JAN-1996  (first entry)
XX
DE   peptide-ligand conjugate #1.
XX
KW   Peptide-ligand; metal-labelled imaging agent; maleimide; metal atom;
KW   technetium; tissue; cell type; organ.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= Pic-Ser
XX
PN   WO9513832-A1.
XX
PD   26-MAY-1995.
XX
PF   16-NOV-1994;   94WO-CA000637.
XX
PR   16-NOV-1993;   93US-00152680.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Pollak A,  Dunn-Dufault R;
XX
DR   WPI; 1995-200205/26.
```

```
XX
PT   Compsns. for generating metal labelled imaging agents - comprising a
PT   solid support, a linking gp. and a ligand cleavable from the linking gp.
PT   by a metal.
XX
PS   Claim 6; Page 21; 31pp; English.
XX
CC   Peptides AAR76218-9 are examples of peptide-ligands used in a method to
CC   produce a metal-labelled imaging agent. The method comprises attaching a
CC   ligand-target molecule e.g. a peptide to a solid surface via a linking
CC   group e.g. maleimide. The ligand can incorporate a metal atom e.g
CC   technetium, which cleaves the linker-ligand bond thus releasing a
CC   labelled peptide. The labelled peptide can then be used for imaging in
CC   tissues, cell types or organs
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          4 TKPPR 8


RESULT 25
AAW11058
ID   AAW11058 standard; peptide; 8 AA.
XX
AC   AAW11058;
XX
DT   03-JUN-1997  (first entry)
XX
DE   Leukocyte-targetted peptide derived radionuclide chelator.
XX
KW   Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
KW   diagnostic imaging.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Sarcosine-Ser"
FT   Modified-site   2
FT                   /note= "protected with Acm group"
XX
PN   WO9603427-A1.
XX
PD   08-FEB-1996.
XX
PF   28-APR-1995;   95WO-CA000249.
XX
PR   22-JUL-1994;   94US-00279155.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Pollak A,  Goodbody A;
XX
DR   WPI; 1996-116994/12.
XX
PT   New peptide derived radionuclide chelators and metal complexes - useful
PT   for diagnostic imaging.
XX
PS   Example 3; Page 14; 30pp; English.
XX
CC   AAW11054-W11059 are peptide derived radionuclide chelators that are
CC   targetted to leukocytes, via the GTKPPR sequence. When the chelators are
CC   coupled to a targeting molecule and labelled with a diagnostically useful
CC   metal, they can be used to detect pathological conditions by diagnostic
CC   imaging. The leukocyte-targetted chelators are useful for the rapid
CC   imaging of sites of local inflammation. Radionuclides used include 99mTc,
CC   64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 186Re, 188Re, 198Au, 199Au, 203Pb, 212Pb
CC   and 212Bi. The coupling of a targeting agent and radionuclide using a
CC   chelating agent is an alternative to the direct labelling of targeting
CC   agents in which radionuclides are typically bound at the more numerous
CC   low-affinity sites, forming unstable complexes. The new conjugates give
CC   better scintigraphic images in rat inflammation studies than known
CC   imaging agents Ga-67, 99mTc-IgG, 111In-WBC and 99mTc-Nanocoll. They image
CC   more rapidly than the known agents and show superior biodistribution
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          4 TKPPR 8


RESULT 26
AAY49844
ID   AAY49844 standard; peptide; 8 AA.
XX
AC   AAY49844;
XX
DT   20-JAN-2000  (first entry)
XX
DE   Tuftsin receptor antagonist chelate conjugate #5.
XX
KW   Tuftsin receptor antagonist; chelate conjugate; radiopharmaceutical;
KW   diagnosis; infection; inflammation; imaging; cancer; tumour.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "6-(hydrazino)-3-pyridylcarbonyl"
FT   Modified-site   2
FT                   /note= "modification on the COOH side group: -Acp-Thr-Lys
FT                   -Pro-Pro-Arg"
FT   Modified-site   3
FT                   /label= Acp
FT                   /note= "6-aminocaproic acid"
XX
PN   WO9951628-A1.
XX
PD   14-OCT-1999.
XX
PF   29-MAR-1999;   99WO-US006824.
XX
PR   03-APR-1998;   98US-0080672P.
XX
PA   (DUPO ) DU PONT PHARM CO.
XX
PI   Edwards DS,  Rajopadhye M;
XX
DR   WPI; 1999-633729/54.
XX
PT   New polypeptides and radiopharmacueticals used for imaging infection,
PT   inflammation and cancer.
XX
PS   Claim 4; Page 71; 80pp; English.
XX
CC   The present invention describes polypeptide compounds of formula (I),
CC   capable of direct transformation into a radiopharmaceutical: Ch-Ln-(X1-X2
CC   -X3-X4-X5)d (I), where X1-X5 = amino acids; Ln = a linking group; Ch = a
CC   metal bonding unit; and d is selected from 1, 2 and 3. The
CC   radiopharmaceuticals are useful for the diagnosis of infection,
CC   inflammation and cancer. The radiopharmaceuticals bind in vivo to the
CC   tuftsin receptor on the surface of white cells which accumulate at the
CC   site of infection and inflammation and can then be detected using
CC   radiation detecting probes or by imaging using a planar or ring gamma
CC   camera. The radiopharmaceuticals can also be used in treating cancer. The
```

```
CC   present sequence represents a specifically claimed tuftsin receptor
CC   antagonist chelate conjugate from the present invention
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          4 TKPPR 8


RESULT 27
ABB08448
ID   ABB08448 standard; peptide; 8 AA.
XX
AC   ABB08448;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Tuftsin receptor antagonist (TKPPR) derivative tetramer (BRU-346).
XX
KW   Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
KW   angiogenesis; tumour cell; cytostatic; antagonist.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   6
FT                   /note= "residue modified by the addition of the following
FT                   block of residues (RPPKT)"
FT   Modified-site   7
FT                   /note= "residue modified by the addition of the following
FT                   block of residues [K(RPPKT)TKPPR]"
FT   Modified-site   8
FT                   /note= "residue modified by the addition of Z-OG,
FT                   (Z=TTDA=4,7,10-Trioxa-1,13-TridecaneDiamine, OG= Oregon
FT                   green)"
XX
PN   WO200191805-A2.
XX
PD   06-DEC-2001.
XX
PF   04-JUN-2001; 2001WO-US018053.
XX
PR   02-JUN-2000; 2000US-00585364.
XX
PA   (BRAC ) BRACCO RES USA.
XX
PI   Von Wronski MA,  Marinelli ER,  Nunn AD,  Pillai R,  Ramalingam K;
PI   Tweedle MF,  Linder K,  Nanjappan P,  Raju N;
XX
DR   WPI; 2002-195523/25.
XX
PT   Composition for use in targeting endothelial cells, tumor cells or other
PT   cells which express NP-1 comprises a compound containing a polypeptide,
PT   linker and substrate.
XX
PS   Example 29; Page 114; 146pp; English.
XX
CC   The invention relates to a composition for use in targeting endothelial
CC   cells, tumour cells, or other cells which express NP-1. The activity of
CC   compositions of the invention may be described as cytostatic. Compounds
CC   of the invention are useful in pharmaceutical compositions for inhibiting
CC   angiogenesis, for imaging and targeting an angiogenic site, endothelial
CC   cells, tumour cells or other cells that express NP-1 in a human or
CC   animal. They may also be used as ultrasound contrast agents, for staging
CC   a tumour in a human or animal, for screening for the ability of an agent
CC   to target endothelial cells, tumour cells or other cells that express NP-
CC   1. They may be used for therapeutic delivery in vivo of a bioactive agent
CC   or for treating an individual exhibiting effects of an angiogenesis or a
CC   related disorder. They may be used for delivering desired nucleic acids
CC   to endothelial cells, tumour cells or other cells expressing NP-1, for
CC   enhancing endothelial or tumour cell-targeted gene therapy, or gene
CC   therapy targeting angiogenic cells, and for treating a human or animal
CC   with a tumour or angiogenesis-related disease. The current sequence
CC   represents a tuftsin receptor antagonist (TKPPR) derivative tetramer BRU-
CC   346
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 28
AAR85539
ID   AAR85539 standard; peptide; 9 AA.
XX
AC   AAR85539;
XX
DT   17-APR-1996  (first entry)
XX
DE   Metal chelating peptide conjugated onto tuftsin antagonist.
XX
KW   chelate; chelating; chelator; diagnostic; imaging; therapy; tuftsin.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .4
FT                   /label= metal_chelating_moiety
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "residue of 2-quinolinic acid, 1- or 3-
FT                   isoquinolinic acid, pyrrole-2-carboxylic acid or indole-2
FT                   -carboxylic acid"
FT   Modified-site   3
FT                   /note= "Cys(Acm)"
FT   Domain          5. .9
FT                   /label= tuftsin_antagonist
XX
PN   WO9517419-A1.
XX
PD   29-JUN-1995.
XX
PF   22-DEC-1994;   94WO-CA000718.
XX
PR   22-DEC-1993;   93US-00171737.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Pollak A,  Goodbody A;
XX
DR   WPI; 1995-240607/31.
XX
PT   New peptide analogue metal chelator(s) - used for chelating radioactive
PT   metals for in-vivo diagnostic imaging or for therapy.
XX
PS   Claim 12; Page 26; 34pp; English.
XX
CC   The patent discloses new metal-chelating peptide analogues of formula
CC   Chel-Xaa-Cys-X, in which: Chel is a chelating moiety which is the residue
CC   of a heterocyclic amino acid analogue such as picolinic acid, dipicolinic
CC   acid, chelidamic acid, 2-carboxypyrazine, 2-carboxypyrimidine, 2-
CC   carboxypyrrole, 2-quinolinic acid or 2- or 3-isoquinolinic acid; Xaa is
CC   an amino acid residue; the Cys residue is optionally protected; and X is
CC   OH, alkoxy, an amino acid residue or a targetting moiety (TM). Preferably
CC   Xaa is Gly or Ser; X is Gly-OH or Gly-TM; and TM is a cell receptor
```

CC binding peptide. The molecule can be used to chelate a diagnostically or
CC therapeutically useful metal such as 99m-Tc, and the targetting moiety
CC can be used to direct the molecule to a site of interest within the body
CC for in-vivo diagnostic imaging or for therapy. The present sequence
CC represents specifically claimed examples of the new peptides, in which
CC the targetting peptide is a tuftsin antagonist
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 29; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 5 TKPPR 9

RESULT 29
AAR85535
ID AAR85535 standard; peptide; 9 AA.
XX
AC AAR85535;
XX
DT 17-APR-1996 (first entry)
XX
DE Metal chelating peptide conjugated onto tuftsin antagonist.
XX
KW chelate; chelating; chelator; diagnostic; imaging; therapy; tuftsin.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Domain 1. .4
FT /label= metal_chelating_moiety
FT Modified-site 1
FT /label= OTHER
FT /note= "picolinic acid residue"
FT Modified-site 3
FT /note= "Cys(Acm)"
FT Domain 5. .9
FT /label= tuftsin_antagonist
XX
PN WO9517419-A1.
XX
PD 29-JUN-1995.
XX
PF 22-DEC-1994; 94WO-CA000718.
XX
PR 22-DEC-1993; 93US-00171737.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Pollak A, Goodbody A;
XX
DR WPI; 1995-240607/31.
XX
PT New peptide analogue metal chelator(s) - used for chelating radioactive
PT metals for in-vivo diagnostic imaging or for therapy.
XX
PS Claim 11; Page 26; 34pp; English.
XX
CC The patent discloses new metal-chelating peptide analogues of formula
CC Chel-Xaa-Cys-X, in which: Chel is a chelating moiety which is the residue
CC of a heterocyclic amino acid analogue such as picolinic acid, dipicolinic
CC acid, chelidamic acid, 2-carboxypyrazine, 2-carboxypyrimidine, 2-
CC carboxypyrrole, 2-quinolinic acid or 2- or 3-isoquinolinic acid; Xaa is
CC an amino acid residue; the Cys residue is optionally protected; and X is
CC OH, alkoxy, an amino acid residue or a targetting moiety (TM). Preferably
CC Xaa is Gly or Ser; X is Gly-OH or Gly-TM; and TM is a cell receptor
CC binding peptide. The molecule can be used to chelate a diagnostically or
CC therapeutically useful metal such as 99m-Tc, and the targetting moiety
CC can be used to direct the molecule to a site of interest within the body
CC for in-vivo diagnostic imaging or for therapy. The present sequence
CC represents a specifically claimed example of the new peptides, in which
CC the targetting peptide is a tuftsin antagonist
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 29; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 5 TKPPR 9

RESULT 30
AAR88735
ID AAR88735 standard; peptide; 9 AA.
XX
AC AAR88735;
XX
DT 10-APR-1996 (first entry)
XX
DE Tuftsin antagonist peptide-metal chelator conjugate.
XX
KW Peptide-chelator conjugate; metal chelator; diagnostic imaging;
KW inflammation; radionuclide; tuftsin; analogue; antagonist.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1. .3
FT /label= chelator
FT /note= "pref. chelates a radionuclide"
FT Misc-difference 1
FT /label= OTHER
FT /note= "picolinic acid or N',N-dimethylglycine"
FT Modified-site 3
FT /label= OTHER
FT /note= "Cys(Acm)"
FT Region 4
FT /label= linking_group
FT Peptide 5. .9
FT /label= tuftsin_antagonist
XX
PN WO9522996-A2.
XX
PD 31-AUG-1995.
XX
PF 24-FEB-1995; 95WO-CA000106.
XX
PR 25-FEB-1994; 94US-00202178.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Goodbody A, Pollak A;
XX
DR WPI; 1995-311386/40.
XX
PT New peptide-chelator conjugate and complex with traceable metal - used to
PT image sites of inflammation in vivo without significant accumulation on
PT the gastrointestinal tract.
XX
PS Claim 16; Page 21; 23pp; English.
XX
CC The present sequence is that of a specifically claimed peptide- chelator
CC conjugate in which a tuftsin antagonist peptide is coupled to a metal
CC chelator, via a linking group. The chelator serves as a labelling site
CC for radionuclide metals such as technetium-99m. The tuftsin antagonist
CC targets the conjugate to macrophages and neutrophils at sites of
CC inflammation without significant accumulation in the gastrointestinal
CC tract (unlike the native tuftsin tetrapeptide). The conjugate is thus
CC useful for diagnostic imaging of inflammation sites, providing an

```
CC   improved target to background ratio
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 31
AAR88738
ID   AAR88738 standard; peptide; 9 AA.
XX
AC   AAR88738;
XX
DT   10-APR-1996  (first entry)
XX
DE   Tuftsin antagonist peptide-metal chelator conjugate.
XX
KW   Peptide-chelator conjugate; metal chelator; diagnostic imaging;
KW   inflammation; radionuclide; tuftsin; analogue; antagonist.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .3
FT                   /label= chelator
FT                   /note= "pref. chelates a radionuclide"
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "N',N-dimethylglycine"
FT   Modified-site   3
FT                   /label= OTHER
FT                   /note= "Cys(Acm)"
FT   Modified-site   4
FT                   /label= bAla
FT                   /note= "linking group"
FT   Peptide         5. .9
FT                   /label= tuftsin_antagonist
XX
PN   WO9522996-A2.
XX
PD   31-AUG-1995.
XX
PF   24-FEB-1995;   95WO-CA000106.
XX
PR   25-FEB-1994;   94US-00202178.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Goodbody A,  Pollak A;
XX
DR   WPI; 1995-311386/40.
XX
PT   New peptide-chelator conjugate and complex with traceable metal - used to
PT   image sites of inflammation in vivo without significant accumulation on
PT   the gastrointestinal tract.
XX
PS   Claim 16; Page 21; 23pp; English.
XX
CC   The present sequence is that of a specifically claimed peptide- chelator
CC   conjugate in which a tuftsin antagonist peptide is coupled to a metal
CC   chelator, via a linking group. The chelator serves as a labelling site
CC   for radionuclide metals such as technetium-99m. The tuftsin antagonist
CC   targets the conjugate to macrophages and neutrophils at sites of
CC   inflammation without significant accumulation in the gastrointestinal
CC   tract (unlike the native tuftsin tetrapeptide). The conjugate is thus
CC   useful for diagnostic imaging of inflammation sites, providing an
CC   improved target to background ratio
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 32
AAR88741
ID   AAR88741 standard; peptide; 9 AA.
XX
AC   AAR88741;
XX
DT   10-APR-1996  (first entry)
XX
DE   Tuftsin antagonist peptide-metal chelator conjugate.
XX
KW   Peptide-chelator conjugate; metal chelator; diagnostic imaging;
KW   inflammation; radionuclide; tuftsin; analogue; antagonist.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .5
FT                   /label= tuftsin_antagonist
FT   Region          6. .7
FT                   /label= linking_group
FT   Modified-site   6
FT                   /label= bAla
FT   Modified-site   7
FT                   /label= bAla
FT   Modified-site   8
FT                   /label= OTHER
FT                   /note= "the side-chain (epsilon) amino group of Lys at
FT                   position 8 forms a peptide bond with the C-terminus of
FT                   the tetrapeptide N',N-dimethylglycyl-Ser-Cys(Acm)-Gly-;
FT                   this sidechain group consists of a tripeptide metal
FT                   chelator and a linking residue (Gly)"
XX
PN   WO9522996-A2.
XX
PD   31-AUG-1995.
XX
PF   24-FEB-1995;   95WO-CA000106.
XX
PR   25-FEB-1994;   94US-00202178.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Goodbody A,  Pollak A;
XX
DR   WPI; 1995-311386/40.
XX
PT   New peptide-chelator conjugate and complex with traceable metal - used to
PT   image sites of inflammation in vivo without significant accumulation on
PT   the gastrointestinal tract.
XX
PS   Claim 19; Page 21; 23pp; English.
XX
CC   The present sequence is that of a specifically claimed peptide- chelator
CC   conjugate in which a tuftsin antagonist peptide is coupled to a metal
CC   chelator, via a linking group. The chelator serves as a labelling site
CC   for radionuclide metals such as technetium-99m. The tuftsin antagonist
CC   targets the conjugate to macrophages and neutrophils at sites of
CC   inflammation without significant accumulation in the gastrointestinal
CC   tract (unlike the native tuftsin tetrapeptide). The conjugate is thus
CC   useful for diagnostic imaging of inflammation sites, providing an
CC   improved target to background ratio
```

```
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 33
AAR76219
ID   AAR76219 standard; peptide; 9 AA.
XX
AC   AAR76219;
XX
DT   12-JAN-1996  (first entry)
XX
DE   Peptide-ligand conjugate #2.
XX
KW   Peptide-ligand; metal-labelled imaging agent; maleimide; metal atom;
KW   technetium; tissue; cell type; organ.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= N,N'-dimethyl-Gly
XX
PN   WO9513832-A1.
XX
PD   26-MAY-1995.
XX
PF   16-NOV-1994;   94WO-CA000637.
XX
PR   16-NOV-1993;   93US-00152680.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Pollak A,  Dunn-Dufault R;
XX
DR   WPI; 1995-200205/26.
XX
PT   Compsns. for generating metal labelled imaging agents - comprising a
PT   solid support, a linking gp. and a ligand cleavable from the linking gp.
PT   by a metal.
XX
PS   Claim 6; Page 21; 31pp; English.
XX
CC   Peptides AAR76218-9 are examples of peptide-ligands used in a method to
CC   produce a metal-labelled imaging agent. The method comprises attaching a
CC   ligand-target molecule e.g. a peptide to a solid surface via a linking
CC   group e.g. maleimide. The ligand can incorporate a metal atom e.g
CC   technetium, which cleaves the linker-ligand bond thus releasing a
CC   labelled peptide. The labelled peptide can then be used for imaging in
CC   tissues, cell types or organs
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 34
AAW11055
ID   AAW11055 standard; peptide; 9 AA.
XX
AC   AAW11055;
XX
DT   03-JUN-1997  (first entry)
XX
DE   Leukocyte-targetted peptide derived radionuclide chelator.
XX
KW   Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
KW   diagnostic imaging.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N, N-dimethyl-Gly"
FT   Modified-site   3
FT                   /note= "protected with Acm group"
XX
PN   WO9603427-A1.
XX
PD   08-FEB-1996.
XX
PF   28-APR-1995;   95WO-CA000249.
XX
PR   22-JUL-1994;   94US-00279155.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Pollak A,  Goodbody A;
XX
DR   WPI; 1996-116994/12.
XX
PT   New peptide derived radionuclide chelators and metal complexes - useful
PT   for diagnostic imaging.
XX
PS   Claim 28; Page 22; 30pp; English.
XX
CC   AAW11054-W11059 are peptide derived radionuclide chelators that are
CC   targetted to leukocytes, via the GTKPPR sequence. When the chelators are
CC   coupled to a targeting molecule and labelled with a diagnostically useful
CC   metal, they can be used to detect pathological conditions by diagnostic
CC   imaging. The leukocyte-targetted chelators are useful for the rapid
CC   imaging of sites of local inflammation. Radionuclides used include 99mTc,
CC   64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 186Re, 188Re, 198Au, 199Au, 203Pb, 212Pb
CC   and 212Bi. The coupling of a targeting agent and radionuclide using a
CC   chelating agent is an alternative to the direct labelling of targeting
CC   agents in which radionuclides are typically bound at the more numerous
CC   low-affinity sites, forming unstable complexes. The new conjugates give
CC   better scintigraphic images in rat inflammation studies than known
CC   imaging agents Ga-67, 99mTc-IgG, 111In-WBC and 99mTc-Nanocoll. They image
CC   more rapidly than the known agents and show superior biodistribution
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 35
AAW11059
ID   AAW11059 standard; peptide; 9 AA.
XX
AC   AAW11059;
XX
DT   03-JUN-1997  (first entry)
XX
DE   Leukocyte-targetted peptide derived radionuclide chelator.
XX
```

```
KW   Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
KW   diagnostic imaging.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Sarcosine-Gly"
FT   Modified-site   3
FT                   /note= "protected with Acm group"
XX
PN   WO9603427-A1.
XX
PD   08-FEB-1996.
XX
PF   28-APR-1995;   95WO-CA000249.
XX
PR   22-JUL-1994;   94US-00279155.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Pollak A,  Goodbody A;
XX
DR   WPI; 1996-116994/12.
XX
PT   New peptide derived radionuclide chelators and metal complexes - useful
PT   for diagnostic imaging.
XX
PS   Example 1; Page 12; 30pp; English.
XX
CC   AAW11054-W11059 are peptide derived radionuclide chelators that are
CC   targetted to leukocytes, via the GTKPPR sequence. When the chelators are
CC   coupled to a targeting molecule and labelled with a diagnostically useful
CC   metal, they can be used to detect pathological conditions by diagnostic
CC   imaging. The leukocyte-targetted chelators are useful for the rapid
CC   imaging of sites of local inflammation. Radionuclides used include 99mTc,
CC   64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 186Re, 188Re, 198Au, 199Au, 203Pb, 212Pb
CC   and 212Bi. The coupling of a targeting agent and radionuclide using a
CC   chelating agent is an alternative to the direct labelling of targeting
CC   agents in which radionuclides are typically bound at the more numerous
CC   low-affinity sites, forming unstable complexes. The new conjugates give
CC   better scintigraphic images in rat inflammation studies than known
CC   imaging agents Ga-67, 99mTc-IgG, 111In-WBC and 99mTc-Nanocoll. They image
CC   more rapidly than the known agents and show superior biodistribution
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 36
AAW11054
ID   AAW11054 standard; peptide; 9 AA.
XX
AC   AAW11054;
XX
DT   03-JUN-1997  (first entry)
XX
DE   Leukocyte-targetted peptide derived radionuclide chelator.
XX
KW   Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
KW   diagnostic imaging.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N, N-dimethyl-Gly"
FT   Modified-site   3
FT                   /note= "protected with Acm group"
XX
PN   WO9603427-A1.
XX
PD   08-FEB-1996.
XX
PF   28-APR-1995;   95WO-CA000249.
XX
PR   22-JUL-1994;   94US-00279155.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Pollak A,  Goodbody A;
XX
DR   WPI; 1996-116994/12.
XX
PT   New peptide derived radionuclide chelators and metal complexes - useful
PT   for diagnostic imaging.
XX
PS   Claim 28; Page 22; 30pp; English.
XX
CC   AAW11054-W11059 are peptide derived radionuclide chelators that are
CC   targetted to leukocytes, via the GTKPPR sequence. When the chelators are
CC   coupled to a targeting molecule and labelled with a diagnostically useful
CC   metal, they can be used to detect pathological conditions by diagnostic
CC   imaging. The leukocyte-targetted chelators are useful for the rapid
CC   imaging of sites of local inflammation. Radionuclides used include 99mTc,
CC   64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 186Re, 188Re, 198Au, 199Au, 203Pb, 212Pb
CC   and 212Bi. The coupling of a targeting agent and radionuclide using a
CC   chelating agent is an alternative to the direct labelling of targeting
CC   agents in which radionuclides are typically bound at the more numerous
CC   low-affinity sites, forming unstable complexes. The new conjugates give
CC   better scintigraphic images in rat inflammation studies than known
CC   imaging agents Ga-67, 99mTc-IgG, 111In-WBC and 99mTc-Nanocoll. They image
CC   more rapidly than the known agents and show superior biodistribution
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 37
AAW11056
ID   AAW11056 standard; peptide; 9 AA.
XX
AC   AAW11056;
XX
DT   03-JUN-1997  (first entry)
XX
DE   Leukocyte-targetted peptide derived radionuclide chelator.
XX
KW   Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
KW   diagnostic imaging.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N, N-diethyl-Gly"
FT   Modified-site   3
FT                   /note= "protected with Acm group"
XX
PN   WO9603427-A1.
XX
PD   08-FEB-1996.
```

XX
PF 28-APR-1995; 95WO-CA000249.
XX
PR 22-JUL-1994; 94US-00279155.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Pollak A, Goodbody A;
XX
DR WPI; 1996-116994/12.
XX
PT New peptide derived radionuclide chelators and metal complexes - useful
PT for diagnostic imaging.
XX
PS Claim 28; Page 22; 30pp; English.
XX
CC AAW11054-W11059 are peptide derived radionuclide chelators that are
CC targetted to leukocytes, via the GTKPPR sequence. When the chelators are
CC coupled to a targeting molecule and labelled with a diagnostically useful
CC metal, they can be used to detect pathological conditions by diagnostic
CC imaging. The leukocyte-targetted chelators are useful for the rapid
CC imaging of sites of local inflammation. Radionuclides used include 99mTc,
CC 64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 186Re, 188Re, 198Au, 199Au, 203Pb, 212Pb
CC and 212Bi. The coupling of a targeting agent and radionuclide using a
CC chelating agent is an alternative to the direct labelling of targeting
CC agents in which radionuclides are typically bound at the more numerous
CC low-affinity sites, forming unstable complexes. The new conjugates give
CC better scintigraphic images in rat inflammation studies than known
CC imaging agents Ga-67, 99mTc-IgG, 111In-WBC and 99mTc-Nanocoll. They image
CC more rapidly than the known agents and show superior biodistribution
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 29; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 5 TKPPR 9

RESULT 38
AAW11057
ID AAW11057 standard; peptide; 9 AA.
XX
AC AAW11057;
XX
DT 03-JUN-1997 (first entry)
XX
DE Leukocyte-targetted peptide derived radionuclide chelator.
XX
KW Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
KW diagnostic imaging.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N, N-dibenzyl-Gly"
FT Modified-site 3
FT /note= "protected with Acm group"
XX
PN WO9603427-A1.
XX
PD 08-FEB-1996.
XX
PF 28-APR-1995; 95WO-CA000249.
XX
PR 22-JUL-1994; 94US-00279155.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Pollak A, Goodbody A;
XX
DR WPI; 1996-116994/12.
XX
PT New peptide derived radionuclide chelators and metal complexes - useful
PT for diagnostic imaging.
XX
PS Claim 28; Page 22; 30pp; English.
XX
CC AAW11054-W11059 are peptide derived radionuclide chelators that are
CC targetted to leukocytes, via the GTKPPR sequence. When the chelators are
CC coupled to a targeting molecule and labelled with a diagnostically useful
CC metal, they can be used to detect pathological conditions by diagnostic
CC imaging. The leukocyte-targetted chelators are useful for the rapid
CC imaging of sites of local inflammation. Radionuclides used include 99mTc,
CC 64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 186Re, 188Re, 198Au, 199Au, 203Pb, 212Pb
CC and 212Bi. The coupling of a targeting agent and radionuclide using a
CC chelating agent is an alternative to the direct labelling of targeting
CC agents in which radionuclides are typically bound at the more numerous
CC low-affinity sites, forming unstable complexes. The new conjugates give
CC better scintigraphic images in rat inflammation studies than known
CC imaging agents Ga-67, 99mTc-IgG, 111In-WBC and 99mTc-Nanocoll. They image
CC more rapidly than the known agents and show superior biodistribution
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 29; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 5 TKPPR 9

RESULT 39
AAW03420
ID AAW03420 standard; peptide; 9 AA.
XX
AC AAW03420;
XX
DT 10-OCT-1997 (first entry)
XX
DE Peptide useful as diagnostic imaging agent.
XX
KW chelator; ligand.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1. .3
FT /label= chelator
FT Modified-site 1
FT /note= "N,N-dimethyl-Gly"
FT Modified-site 3
FT /label= ligand_derivatised_residue
FT /note= " a galactosyl residue is attached to the thiol
FT group of this Cys via a cleavable maleimido-containing
FT coupling group"
FT Region 5. .9
FT /label= targetting_molecule
XX
PN WO9638185-A1.
XX
PD 05-DEC-1996.
XX
PF 16-MAY-1996; 96WO-CA000310.
XX
PR 31-MAY-1995; 95US-00454859.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Pollak A;

```
XX
DR   WPI; 1997-042662/04.
XX
PT   New cpd. for use in diagnostic imaging - consists of imaging agent, which
PT   is chelator for traceable metal, coupled by metal-cleavable bond to
PT   ligand which localises at different site.
XX
PS   Claim 10-14; Page 16; 26pp; English.
XX
CC   The patent discloses new peptides useful for diagnostic imaging. The
CC   peptides comprise (1) an imaging agent (a chelator for a traceable metal)
CC   which localises selectively at an in-vivo site of diagnostic interest;
CC   (2) a ligand which localises at a point (in vivo) removed from the site
CC   of diagnostic interest; and (3) a metal-cleavable bond coupling the
CC   chelator of the imaging agent to the ligand, this bond being cleaved upon
CC   formation of a coordination complex of the metal and chelator. Preferably
CC   the chelator is the sequence N,N-dimethyl- Gly-Ser-Cys; the ligand is a
CC   galactosyl residue; the traceable metal is 99m-Tc; and a targetting
CC   molecule (especially TKPPR) is attached to the chelator. Upon labelling,
CC   the ligand is cleaved, leaving the labelled imaging agent free to
CC   localise at the site of diagnostic interest unhindered, while the ligand
CC   and any unlabelled imaging agent are sequestered to the removed site. By
CC   sequestering unlabelled imaging agent, the labelled imaging agent does
CC   not compete to occupy the site of interest, resulting in images of
CC   enhanced resolution. The present sequence is the preferred molecule used
CC   for imaging, in which a galactosyl residue is attached to the thiol group
CC   of Cys(3) via a cleavable maleimido coupling group
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 40
AAY23752
ID   AAY23752 standard; peptide; 9 AA.
XX
AC   AAY23752;
XX
DT   09-SEP-1999  (first entry)
XX
DE   Peptide RP502 used to prepare chirally pure peptides.
XX
KW   Chirally pure peptide; metal complex; diagnostic imaging; radioimaging.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "DimethylGly"
FT   Modified-site   2
FT                   /note= "t-butylGly"
FT   Modified-site   3
FT                   /note= "Acm protecting group attached"
XX
PN   WO9933863-A1.
XX
PD   08-JUL-1999.
XX
PF   23-DEC-1998;   98WO-CA001201.
XX
PR   24-DEC-1997;   97US-00997802.
PR   30-DEC-1997;   97CA-02226226.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Pollak A,  Fauconnier T,  Wong E;
XX
DR   WPI; 1999-419086/35.
XX
PT   Chirally pure peptides for preparation of metal complexes, used for
PT   diagnostic imaging.
XX
PS   Example 1; Page 18; 35pp; English.
XX
CC   The specification describes chirally pure peptides which are used for
CC   preparation of metal complexes. When converted to complexes with
CC   (radioactive) metals or their oxides or nitrides, the peptides are useful
CC   for diagnostic (radio)imaging (by labeling biomolecules). The present
CC   peptide is used in the course of the invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          5 TKPPR 9


RESULT 41
AAR88739
ID   AAR88739 standard; peptide; 10 AA.
XX
AC   AAR88739;
XX
DT   10-APR-1996  (first entry)
XX
DE   Tuftsin antagonist peptide-metal chelator conjugate.
XX
KW   Peptide-chelator conjugate; metal chelator; diagnostic imaging;
KW   inflammation; radionuclide; tuftsin; analogue; antagonist.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .3
FT                   /label= chelator
FT                   /note= "pref. chelates a radionuclide"
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "N',N-dimethylglycine"
FT   Modified-site   3
FT                   /label= OTHER
FT                   /note= "Cys(Acm)"
FT   Region          4. .5
FT                   /label= linking_group
FT   Modified-site   4
FT                   /label= bAla
FT   Modified-site   5
FT                   /label= bAla
FT   Peptide         6. .10
FT                   /label= tuftsin_antagonist
XX
PN   WO9522996-A2.
XX
PD   31-AUG-1995.
XX
PF   24-FEB-1995;   95WO-CA000106.
XX
PR   25-FEB-1994;   94US-00202178.
XX
PA   (RESO-) RESOLUTION PHARM INC.
XX
PI   Goodbody A,  Pollak A;
XX
DR   WPI; 1995-311386/40.
XX
```

PT New peptide-chelator conjugate and complex with traceable metal - used to
PT image sites of inflammation in vivo without significant accumulation on
PT the gastrointestinal tract.
XX
PS Claim 16; Page 21; 23pp; English.
XX
CC The present sequence is that of a specifically claimed peptide- chelator
CC conjugate in which a tuftsin antagonist peptide is coupled to a metal
CC chelator, via a linking group. The chelator serves as a labelling site
CC for radionuclide metals such as technetium-99m. The tuftsin antagonist
CC targets the conjugate to macrophages and neutrophils at sites of
CC inflammation without significant accumulation in the gastrointestinal
CC tract (unlike the native tuftsin tetrapeptide). The conjugate is thus
CC useful for diagnostic imaging of inflammation sites, providing an
CC improved target to background ratio
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||||
Db 6 TKPPR 10

RESULT 42
ADE81165
ID ADE81165 standard; peptide; 10 AA.
XX
AC ADE81165;
XX
DT 29-JAN-2004 (first entry)
XX
DE Tufstin metallopeptide analogue #48.
XX
KW immunostimulant; immunosuppressive; tuftsin receptor binder;
KW tuftsin receptor; tufstin receptor analogue; imaging; infection;
KW inflammation; immunostimulatory; immune system disorder; analgesic;
KW CNS condition; whole body imaging; radiotherapy;
KW tuftsin metallopeptide analogue.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 5. .6
FT /note= "Residues joined by NH-(CH2)6-CO"
FT Misc-difference 7
FT /note= "D-form residue"
FT Misc-difference 9
FT /note= "D-form residue"
XX
PN US2003059422-A1.
XX
PD 27-MAR-2003.
XX
PF 30-AUG-1999; 99US-00387715.
XX
PR 07-JUN-1995; 95US-00476652.
PR 05-JUN-1996; 96US-00660697.
PR 18-MAR-1998; 98US-0078373P.
PR 14-DEC-1998; 98US-0112235P.
PR 18-MAR-1999; 99WO-US005693.
XX
PA (SHAR/) SHARMA S D.
XX
PI Sharma SD;
XX
DR WPI; 2003-596563/56.
XX
PT Tuftsin metallopeptide, useful e.g. for imaging site of infection or
PT inflammation, comprises metal ion-binding backbone including at least two
PT contiguous amino acids for complexing with metal ion.
XX
PS Claim 17; Page 11; 12pp; English.
XX
CC The invention describes a peptide or its salt comprising a metal ion-
CC binding backbone including at least two contiguous amino acids for
CC complexing with the metal ion. The peptide is specific for the tuftsin
CC receptor on complexing the metal ion-binding backbone with the metal ion.
CC The peptide is useful for imaging a site of infection or inflammation;
CC for causing an immunostimulatory response in mammals; for treating immune
CC system disorders; in biological, pharmaceutical and radiopharmaceutical
CC applications; as an analgesic in the treatment of CNS conditions; and in
CC whole body imaging and radiotherapy. The peptide complexed with the metal
CC ion is resistant to enzymatic degradation. The affinity of the peptide
CC for the tuftsin receptor is higher when the metal ion-binding backbone is
CC complexed with the metal ion than that when the metal ion-binding
CC backbone is not complexed with the metal ion. The Tc-labeled Thr-D-Lys-
CC Gly-D-Cys-Arg is the most potent existing tuftsin molecule. The peptide-
CC metal ion complexes have a higher level of stability, and are less
CC susceptible to proteolysis than the uncomplexed or existing peptides. The
CC peptide analogue is not conformationally restricted in the absence of a
CC metal ion but has high potency and concomitant conformational restriction
CC on complexation with a metal ion. The metal complexation in the peptide
CC causes specific regional conformational restrictions in the peptide, so
CC that the peptide conformation at the metal binding site is
CC conformationally fixed on metal complexation. The complexation of the
CC peptide to a metal ion alters the in vivo distribution profile, rate and
CC mode of clearance from the body, bioavailability and pharmacokinetics in
CC mammals. The peptide metal ion complex stimulates polymorphonuclear
CC granulocytes, monocytes and macrophages towards phagocytosis, produces
CC higher titer antibodies, and can transit the gut-blood barrier without
CC significant enzymatic or peptidase degradation. This is the amino acid
CC sequence of a tuftsin metallopeptide analogue of the invention.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||||
Db 1 TKPPR 5

RESULT 43
ADE81166
ID ADE81166 standard; peptide; 10 AA.
XX
AC ADE81166;
XX
DT 29-JAN-2004 (first entry)
XX
DE Tufstin metallopeptide analogue #49.
XX
KW immunostimulant; immunosuppressive; tuftsin receptor binder;
KW tuftsin receptor; tufstin receptor analogue; imaging; infection;
KW inflammation; immunostimulatory; immune system disorder; analgesic;
KW CNS condition; whole body imaging; radiotherapy;
KW tuftsin metallopeptide analogue.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "D-form residue"
FT Misc-difference 4
FT /note= "D-form residue"
FT Modified-site 5. .6
FT /note= "Residues joined by NH-(CH2)6-CO"
XX
PN US2003059422-A1.
XX

PD 27-MAR-2003.
XX
PF 30-AUG-1999; 99US-00387715.
XX
PR 07-JUN-1995; 95US-00476652.
PR 05-JUN-1996; 96US-00660697.
PR 18-MAR-1998; 98US-0078373P.
PR 14-DEC-1998; 98US-0112235P.
PR 18-MAR-1999; 99WO-US005693.
XX
PA (SHAR/) SHARMA S D.
XX
PI Sharma SD;
XX
DR WPI; 2003-596563/56.
XX
PT Tuftsin metallopeptide, useful e.g. for imaging site of infection or
PT inflammation, comprises metal ion-binding backbone including at least two
PT contiguous amino acids for complexing with metal ion.
XX
PS Claim 17; Page 11; 12pp; English.
XX
CC The invention describes a peptide or its salt comprising a metal ion-
CC binding backbone including at least two contiguous amino acids for
CC complexing with the metal ion. The peptide is specific for the tuftsin
CC receptor on complexing the metal ion-binding backbone with the metal ion.
CC The peptide is useful for imaging a site of infection or inflammation;
CC for causing an immunostimulatory response in mammals; for treating immune
CC system disorders; in biological, pharmaceutical and radiopharmaceutical
CC applications; as an analgesic in the treatment of CNS conditions; and in
CC whole body imaging and radiotherapy. The peptide complexed with the metal
CC ion is resistant to enzymatic degradation. The affinity of the peptide
CC for the tuftsin receptor is higher when the metal ion-binding backbone is
CC complexed with the metal ion than that when the metal ion-binding
CC backbone is not complexed with the metal ion. The Tc-labeled Thr-D-Lys-
CC Gly-D-Cys-Arg is the most potent existing tuftsin molecule. The peptide-
CC metal ion complexes have a higher level of stability, and are less
CC susceptible to proteolysis than the uncomplexed or existing peptides. The
CC peptide analogue is not conformationally restricted in the absence of a
CC metal ion but has high potency and concomitant conformational restriction
CC on complexation with a metal ion. The metal complexation in the peptide
CC causes specific regional conformational restrictions in the peptide, so
CC that the peptide conformation at the metal binding site is
CC conformationally fixed on metal complexation. The complexation of the
CC peptide to a metal ion alters the in vivo distribution profile, rate and
CC mode of clearance from the body, bioavailability and pharmacokinetics in
CC mammals. The peptide metal ion complex stimulates polymorphonuclear
CC granulocytes, monocytes and macrophages towards phagocytosis, produces
CC higher titer antibodies, and can transit the gut-blood barrier without
CC significant enzymatic or peptidase degradation. This is the amino acid
CC sequence of a tuftsin metallopeptide analogue of the invention.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||| |
Db 6 TKPPR 10

RESULT 44
AAR88736
ID AAR88736 standard; peptide; 11 AA.
XX
AC AAR88736;
XX
DT 10-APR-1996 (first entry)
XX
DE Tuftsin antagonist peptide-metal chelator conjugate.
XX
KW Peptide-chelator conjugate; metal chelator; diagnostic imaging;
KW inflammation; radionuclide; tuftsin; analogue; antagonist.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1. .3
FT /label= chelator
FT /note= "pref. chelates a radionuclide"
FT Misc-difference 1
FT /label= OTHER
FT /note= "picolinic acid or benzoyl-mercaptoacetic acid"
FT Modified-site 3
FT /label= OTHER
FT /note= "Cys(Acm)"
FT Region 4. .6
FT /label= linking_group
FT Peptide 7. .11
FT /label= tuftsin_antagonist
XX
PN WO9522996-A2.
XX
PD 31-AUG-1995.
XX
PF 24-FEB-1995; 95WO-CA000106.
XX
PR 25-FEB-1994; 94US-00202178.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Goodbody A, Pollak A;
XX
DR WPI; 1995-311386/40.
XX
PT New peptide-chelator conjugate and complex with traceable metal - used to
PT image sites of inflammation in vivo without significant accumulation on
PT the gastrointestinal tract.
XX
PS Claim 16; Page 21; 23pp; English.
XX
CC The present sequence is that of a specifically claimed peptide- chelator
CC conjugate in which a tuftsin antagonist peptide is coupled to a metal
CC chelator, via a linking group. The chelator serves as a labelling site
CC for radionuclide metals such as technetium-99m. The tuftsin antagonist
CC targets the conjugate to macrophages and neutrophils at sites of
CC inflammation without significant accumulation in the gastrointestinal
CC tract (unlike the native tuftsin tetrapeptide). The conjugate is thus
CC useful for diagnostic imaging of inflammation sites, providing an
CC improved target to background ratio
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 29; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||| |
Db 7 TKPPR 11

RESULT 45
ADD10691
ID ADD10691 standard; peptide; 11 AA.
XX
AC ADD10691;
XX
DT 01-JAN-2004 (first entry)
XX
DE Tuftsin analogue peptide tetramer.
XX
KW Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
KW antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;

```
KW   visualisation therapy; radiotherapy.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   6
FT                   /label= OTHER
FT                   /note= "Lys is covalently linked to a further TKPPRKRPPKT
FT                   peptide whose  Lys (6) is linked to Gly-TTDA (4,7,10-
FT                   Trioxa1,13-TridecaneDiamine)-oregon green"
XX
PN   US2002147136-A1.
XX
PD   10-OCT-2002.
XX
PF   04-JUN-2001; 2001US-00871974.
XX
PR   02-JUN-2000; 2000US-00585364.
XX
PA   (VWRO/) VON WRONSKI M A.
PA   (MARI/) MARINELLI E R.
PA   (NUNN/) NUNN A D.
PA   (PILL/) PILLAI R.
PA   (RAMA/) RAMALINGAM K.
PA   (TWEE/) TWEEDLE M F.
PA   (LIND/) LINDER K.
PA   (NANJ/) NANJAPPAN P.
PA   (RAJU/) RAJU N.
XX
PI   Von Wronski MA,  Marinelli ER,  Nunn AD,  Pillai R,  Ramalingam K;
PI   Tweedle MF,  Linder K,  Nanjappan P,  Raju N;
XX
DR   WPI; 2003-800817/75.
XX
PT   Composition used in targeting endothelial cells e.g. tumor cells
PT   comprises compounds containing monomers, multimers or polymers of L-
PT   arginine-L-threonyl-L-lysyl-L-prolyl-L-prolyl.
XX
PS   Example 29; Page 66; 85pp; English.
XX
CC   The invention relates to a composition (A1) comprising compounds
CC   containing monomers, multimers or polymers of TKPPR (ADD10684).
CC   Composition (A1) comprises a compound of formula A-L-B 1, where A is the
CC   TKPPR peptide, L is a linker moiety (of formula given in the
CC   specification) and B is a substrate (or a phospholipid group,
CC   derivatisable bead attached to a fluorescent or radioactive marker,
CC   bioactive agent, delivery vehicle for genetic material, drug or
CC   therapeutic, or chelating group (preferably N 4 , S 4 , N 3 S, N 2 S 2 or
CC   NS 3 ) comprising oxa-PnAO complexed with 99m Tc). The compound
CC   specifically binds to NP-1 (Vascular endothelial growth factor binding
CC   receptor transmembrane glycoprotein) or cells that express NP-1 with
CC   avidity of at least that of TKPPR. Also included are an ultrasound
CC   contrast agent (c1) comprising a suspension of gas filled microbubbles
CC   comprising the TKPPR compound, an ultrasound contrast agent (c2)
CC   comprising a suspension of gas filled microballoons comprising the TKPPR
CC   compound, preparation of the TKPPR compound (which comprises conjugating
CC   the monomer, multimer or polymer of TKPPR or its analogue with a linker
CC   to obtain a compound of formula A-L, forming a covalent or non-covalent
CC   bond between A-L and the substrate B 1 or forming a covalent bond between
CC   B 1 and the linker to form a conjugate B-L followed by conjugation with
CC   the monomer), and a kit for preparing a radiopharmaceutical comprising
CC   the compound. The compound used for targeting endothelial cells, tumour
CC   cells or other cells which express NP-1, for inhibiting angiogenesis, for
CC   ultrasound imaging, staging a tumour, screening at least one targeted
CC   ultrasound contrast agent for the ability to target endothelial cells,
CC   tumour cells or other cells which express NP-1, for the therapeutic
CC   delivery in vivo of a bioactive agent and for delivering desired nucleic
CC   acids to endothelial cells, tumour cells or other cells which express NP-
CC   1. The composition is also useful for visualisation therapy or
CC   radiotherapy of endothelial cells. The present sequence is a TKPPR
CC   tetramer peptide.
XX
SQ   Sequence 11 AA;
```

```
  Query Match           100.0%;  Score 29;  DB 7;  Length 11;
  Best Local Similarity  100.0%;  Pred. No. 79;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 46
ADD10690
ID   ADD10690 standard; peptide; 11 AA.
XX
AC   ADD10690;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Tuftsin analogue peptide dimer.
XX
KW   Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
KW   antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
KW   visualisation therapy; radiotherapy.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   6
FT                   /label= OTHER
FT                   /note= "Lys is covalently linked to Gly-TTDA (4,7,10-
FT                   Trioxa1,13-TridecaneDiamine)"
XX
PN   US2002147136-A1.
XX
PD   10-OCT-2002.
XX
PF   04-JUN-2001; 2001US-00871974.
XX
PR   02-JUN-2000; 2000US-00585364.
XX
PA   (VWRO/) VON WRONSKI M A.
PA   (MARI/) MARINELLI E R.
PA   (NUNN/) NUNN A D.
PA   (PILL/) PILLAI R.
PA   (RAMA/) RAMALINGAM K.
PA   (TWEE/) TWEEDLE M F.
PA   (LIND/) LINDER K.
PA   (NANJ/) NANJAPPAN P.
PA   (RAJU/) RAJU N.
XX
PI   Von Wronski MA,  Marinelli ER,  Nunn AD,  Pillai R,  Ramalingam K;
PI   Tweedle MF,  Linder K,  Nanjappan P,  Raju N;
XX
DR   WPI; 2003-800817/75.
XX
PT   Composition used in targeting endothelial cells e.g. tumor cells
PT   comprises compounds containing monomers, multimers or polymers of L-
PT   arginine-L-threonyl-L-lysyl-L-prolyl-L-prolyl.
XX
PS   Example 29; Page 66; 85pp; English.
XX
CC   The invention relates to a composition (A1) comprising compounds
CC   containing monomers, multimers or polymers of TKPPR (ADD10684).
CC   Composition (A1) comprises a compound of formula A-L-B 1, where A is the
CC   TKPPR peptide, L is a linker moiety (of formula given in the
CC   specification) and B is a substrate (or a phospholipid group,
CC   derivatisable bead attached to a fluorescent or radioactive marker,
CC   bioactive agent, delivery vehicle for genetic material, drug or
CC   therapeutic, or chelating group (preferably N 4 , S 4 , N 3 S, N 2 S 2 or
CC   NS 3 ) comprising oxa-PnAO complexed with 99m Tc). The compound
CC   specifically binds to NP-1 (Vascular endothelial growth factor binding
CC   receptor transmembrane glycoprotein) or cells that express NP-1 with
CC   avidity of at least that of TKPPR. Also included are an ultrasound
```

```
CC   contrast agent (c1) comprising a suspension of gas filled microbubbles
CC   comprising the TKPPR compound, an ultrasound contrast agent (c2)
CC   comprising a suspension of gas filled microballoons comprising the TKPPR
CC   compound, preparation of the TKPPR compound (which comprises conjugating
CC   the monomer, multimer or polymer of TKPPR or its analogue with a linker
CC   to obtain a compound of formula A-L, forming a covalent or non-covalent
CC   bond between A-L and the substrate B 1 or forming a covalent bond between
CC   B 1 and the linker to form a conjugate B-L followed by conjugation with
CC   the monomer), and a kit for preparing a radiopharmaceutical comprising
CC   the compound. The compound used for targeting endothelial cells, tumour
CC   cells or other cells which express NP-1, for inhibiting angiogenesis, for
CC   ultrasound imaging, staging a tumour, screening at least one targeted
CC   ultrasound contrast agent for the ability to target endothelial cells,
CC   tumour cells or other cells which express NP-1, for the therapeutic
CC   delivery in vivo of a bioactive agent and for delivering desired nucleic
CC   acids to endothelial cells, tumour cells or other cells which express NP-
CC   1. The composition is also useful for visualisation therapy or
CC   radiotherapy of endothelial cells. The present sequence is a TKPPR dimer
CC   peptide.
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 29;  DB 7;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 79;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          1 TKPPR 5


RESULT 47
ABB08446
ID   ABB08446 standard; peptide; 12 AA.
XX
AC   ABB08446;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Tuftsin receptor antagonist (TKPPR) derivative peptide 3.
XX
KW   Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
KW   angiogenesis; tumour cell; cytostatic; antagonist.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "residue modified by the following; F-108,
FT                   OCH2CONH"
FT   Modified-site   2
FT                   /note= "residue modified by (tBu)"
FT   Modified-site   3
FT                   /note= "residue modified by (Mtt)"
FT   Modified-site   6
FT                   /note= "residue modified by (Pmc)-OtBu"
FT   Modified-site   7
FT                   /note= "residue modified by the following; F-108,
FT                   OCH2CONH"
FT   Modified-site   8
FT                   /note= "residue modified by (tBu)"
FT   Modified-site   9
FT                   /note= "residue modified by (Mtt)"
FT   Modified-site   12
FT                   /note= "residue modified by (Pmc)-OtBu"
XX
PN   WO200191805-A2.
XX
PD   06-DEC-2001.
XX
PF   04-JUN-2001; 2001WO-US018053.
XX
PR   02-JUN-2000; 2000US-00585364.
XX
PA   (BRAC ) BRACCO RES USA.
XX
PI   Von Wronski MA,  Marinelli ER,  Nunn AD,  Pillai R,  Ramalingam K;
PI   Tweedle MF,  Linder K,  Nanjappan P,  Raju N;
XX
DR   WPI; 2002-195523/25.
XX
PT   Composition for use in targeting endothelial cells, tumor cells or other
PT   cells which express NP-1 comprises a compound containing a polypeptide,
PT   linker and substrate.
XX
PS   Example 25; Page 95; 146pp; English.
XX
CC   The invention relates to a composition for use in targeting endothelial
CC   cells, tumour cells, or other cells which express NP-1. The activity of
CC   compositions of the invention may be described as cytostatic. Compounds
CC   of the invention are useful in pharmaceutical compositions for inhibiting
CC   angiogenesis, for imaging and targeting an angiogenic site, endothelial
CC   cells, tumour cells or other cells that express NP-1 in a human or
CC   animal. They may also be used as ultrasound contrast agents, for staging
CC   a tumour in a human or animal, for screening for the ability of an agent
CC   to target endothelial cells, tumour cells or other cells that express NP-
CC   1. They may be used for therapeutic delivery in vivo of a bioactive agent
CC   or for treating an individual exhibiting effects of an angiogenesis or a
CC   related disorder. They may be used for delivering desired nucleic acids
CC   to endothelial cells, tumour cells or other cells expressing NP-1, for
CC   enhancing endothelial or tumour cell-targeted gene therapy, or gene
CC   therapy targeting angiogenic cells, and for treating a human or animal
CC   with a tumour or angiogenesis-related disease. The current sequence
CC   represents a tuftsin receptor antagonist (TKPPR) derivative of the
CC   invention
XX
SQ   Sequence 12 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 85;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TKPPR 5
              |||||
Db          2 TKPPR 6


RESULT 48
ABB08449
ID   ABB08449 standard; peptide; 20 AA.
XX
AC   ABB08449;
XX
DT   01-JUL-2002  (first entry)
XX
DE   99mTc radiopharmaceutical.
XX
KW   Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
KW   angiogenesis; tumour cell; cytostatic; antagonist; radiopharmaceutical.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "residue modified by the addition of 99mTc-Oxa
FT                   PnAO"
XX
PN   WO200191805-A2.
XX
PD   06-DEC-2001.
XX
PF   04-JUN-2001; 2001WO-US018053.
XX
PR   02-JUN-2000; 2000US-00585364.
XX
PA   (BRAC ) BRACCO RES USA.
```

XX
PI Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
PI Tweedle MF, Linder K, Nanjappan P, Raju N;
XX
DR WPI; 2002-195523/25.
XX
PT Composition for use in targeting endothelial cells, tumor cells or other
PT cells which express NP-1 comprises a compound containing a polypeptide,
PT linker and substrate.
XX
PS Example 37; Page 128; 146pp; English.
XX
CC The invention relates to a composition for use in targeting endothelial
CC cells, tumour cells, or other cells which express NP-1. The activity of
CC compositions of the invention may be described as cytostatic. Compounds
CC of the invention are useful in pharmaceutical compositions for inhibiting
CC angiogenesis, for imaging and targeting an angiogenic site, endothelial
CC cells, tumour cells or other cells that express NP-1 in a human or
CC animal. They may also be used as ultrasound contrast agents, for staging
CC a tumour in a human or animal, for screening for the ability of an agent
CC to target endothelial cells, tumour cells or other cells that express NP-
CC 1. They may be used for therapeutic delivery in vivo of a bioactive agent
CC or for treating an individual exhibiting effects of an angiogenesis or a
CC related disorder. They may be used for delivering desired nucleic acids
CC to endothelial cells, tumour cells or other cells expressing NP-1, for
CC enhancing endothelial or tumour cell-targeted gene therapy, or gene
CC therapy targeting angiogenic cells, and for treating a human or animal
CC with a tumour or angiogenesis-related disease. The current sequence
CC represents a tuftsin receptor antagonist (TKPPR)-metal chelator conjugate
CC complexed with a radioactive metal that is described in the specification
CC as 99mTc radiopharmaceutical
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 1 TKPPR 5

RESULT 49
ABP34343
ID ABP34343 standard; protein; 53 AA.
XX
AC ABP34343;
XX
DT 08-JUL-2002 (first entry)
XX
DE Human ORF3316 protein, SEQ ID NO:6632.
XX
KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnerary;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX
OS Homo sapiens.
XX
PN WO200190366-A2.
XX
PD 29-NOV-2001.
XX
PF 24-MAY-2001; 2001WO-US017076.
XX
PR 24-MAY-2000; 2000US-0206690P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Leach MD, Shimkets RA;
XX
DR WPI; 2002-106200/14.
DR N-PSDB; ABN78369.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation.
XX
PS Claim 10; Page 1905; 2508pp; English.
XX
CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antiinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases
XX
SQ Sequence 53 AA;

Query Match 100.0%; Score 29; DB 5; Length 53;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|||||
Db 43 TKPPR 47

RESULT 50
ABG07007
ID ABG07007 standard; protein; 65 AA.
XX
AC ABG07007;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #6998.

XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS71194.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 37366; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 65 AA;

Query Match 100.0%; Score 29; DB 4; Length 65;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
||||| |
Db 27 TKPPR 31

Search completed: March 3, 2004, 12:17:18
Job time : 60 secs